

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**DETECÇÃO DE OSCILAÇÕES EM PROCESSOS INDUSTRIAIS
BASEADA NO ENVELOPE ESPECTRAL**

DISSERTAÇÃO DE MESTRADO

HEITOR GUZZO DE FARIA

VITÓRIA

2018

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**DETECÇÃO DE OSCILAÇÕES EM PROCESSOS INDUSTRIAIS
BASEADA NO ENVELOPE ESPECTRAL**

HEITOR GUZZO DE FARIA

Dissertação de Mestrado apresentada como
requisito parcial para obtenção do título de
Mestre em Engenharia Elétrica.

Orientador:
Prof. Dr. Celso José Munaro.

VITÓRIA
2018

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Setorial Tecnológica,
Universidade Federal do Espírito Santo, ES, Brasil)
Sandra Mara Borges Campos – CRB-6 ES-000593/O

F224d Faria, Heitor Guzzo de, 1990-
Detecção de oscilações em processos industriais baseada no
envelope espectral / Heitor Guzzo de Faria. – 2018.
91 f. : il.

Orientador: Celso José Munaro.
Dissertação (Mestrado em Engenharia Elétrica) –
Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Processos industriais. 2. Sistemas inteligentes de controle.
3. Engenharia elétrica. 4. Detecção de oscilações. 5. Envelope
espectral. 6. Coeficiente de variação. I. Munaro, Celso José. II.
Universidade Federal do Espírito Santo. Centro Tecnológico. III.
Título.

CDU: 621.3

HEITOR GUZZO DE FARIA

**DETECÇÃO DE OSCILAÇÕES EM PROCESSOS INDUSTRIAIS
BASEADA NO ENVELOPE ESPECTRAL**


Dissertação submetida ao programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Mestre em Engenharia Elétrica.

Trabalho aprovado, Vitória, 10 de abril de 2018.

COMISSÃO EXAMINADORA



Prof. Dr. Celso José Munaro
Universidade Federal do Espírito Santo
Orientador



Prof. José Leandro Félix Salles
Universidade Federal do Espírito Santo



Prof. Jorge Otávio Trierweiler
Universidade Federal do Rio Grande do Sul

“Work clears the mind and creates personality.”

(Fyodor Dostoevsky)

“A adversidade é um trampolim para a maturidade.”

(Charles Colton)

Agradecimentos

Agradeço primeiramente a Deus, por ter me dado a oportunidade de ingressar e concluir um curso de mestrado num país onde o acesso à educação é difícil e muitos não têm oportunidade de cursar nível superior, muito menos pós-graduação. Agradeço também a Deus por ter me dado saúde, paciência, persistência, inteligência e resiliência para superar os momentos difíceis e concluir o curso de mestrado.

Gostaria de agradecer também aos meus pais, Wildson e Rosane, a minha irmã Luisa, minhas avós, tias, tios, primas e primos e a toda minha família, por todo tipo de apoio que me deram e pela compreensão nos muitos momentos em que me ausentei devido à dedicação necessária nas minhas atividades do mestrado.

Agradeço também aos meus amigos e colegas do mestrado pela companhia, compreensão, e suporte que deram durante o curso. Em alguns momentos apenas foram pacientes em ouvir minhas ideias e em outros momentos auxiliando diretamente com recomendações, conselhos e esclarecendo dúvidas. Agradecimentos em especial aos colegas do LCI: Vinícius, Gustavo, Arthur, Débora, Felipe, Renata, Eduardo, Ademir, Diego, Uberdan e Marisa.

Agradeço também ao meu orientador Dr. Celso José Munaro pela paciência e pelo suporte durante a orientação da dissertação. Agradecimentos também ao Programa de Pós-Graduação em Engenharia Elétrica (PPGEE) da UFES e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro dado durante os 2 anos de curso e à Petrobrás pela disponibilização de dados para análise.

Resumo

Distúrbios oscilatórios são problemas frequentes em processos industriais. A detecção e diagnóstico de oscilações é de grande importância pois além de indicarem algum defeito ou desgaste de equipamento, o distúrbio pode se propagar por toda a planta provocando aumento da variabilidade e consequentemente redução do desempenho do processo e da qualidade do produto final. Este trabalho apresenta uma proposta de detecção automática de oscilações baseado no envelope espectral. A automatização da metodologia apresentada na literatura possibilita a implementação na indústria sem a necessidade de um especialista operando o algoritmo. Foi realizada a seleção dos segmentos de dados e filtragem para garantir que sinais sem distúrbios fossem utilizados nas análises. Os sinais de menor impacto foram descartados através do coeficiente de variação. O espectro dos sinais restantes foi estimado utilizando a correlação entre os sinais e o envelope foi calculado através do maior autovalor do espectro estimado. A seleção de intervalos de frequência foi realizada de forma automática. Os intervalos de frequência mais relevantes foram selecionados utilizando como critério a energia percentual. A indicação de sinais onde as oscilações estão presentes foi realizada através da verificação do resultado de teste de hipótese. Os sinais em que o valor do teste ultrapassou o limiar da distribuição qui-quadrado com 2 graus de liberdade indicam a existência da oscilação naquele sinal. Estudos de caso foram realizados utilizando dados industriais de uma unidade estacionária de processamento de óleo e gás mostrando que foi possível obter uma boa repetibilidade dos resultados de detecção em análises subsequentes.

Abstract

Oscillatory disturbance is a common issue in industrial processes. The detection and diagnosis of oscillations is important because it can indicate a fault or wear on equipment. The detection is also important because oscillation can propagate throughout the plant causing increase in variability, reduction of process performance and reduction of product quality. This thesis presents a proposal of oscillation detection based on the spectral envelope method. The automation of the methodology presented in the literature allows the implementation in industry without the requirement of a human expert to run the algorithm. Data segmentation was performed together with filtering to guarantee that disturbance free data was used in the analysis. The signals with lower impact were eliminated using the coefficient of variation. The spectrum of the remaining signals was estimated using the correlation between signals and the spectral envelope was calculated via the greatest eigenvalue of the estimated spectrum. Band selection was performed in an automated manner and the most relevant ones were selected using a percentage energy threshold. The oscillation indication was confirmed using a statistical hypothesis test. Signals with values over the threshold of a chi-squared distribution with 2 degrees of freedom indicate the presence of oscillation in the signal. Case studies were carried out using industrial data from an oil and gas processing platform showing that repeatability of detection results was achieved in subsequent analysis.

Lista de Figuras

Figura 1 – Envelope calculado utilizando correlação cruzada e envelope somando espectros	26
Figura 2 – Envelope espectral e correspondente resultado do teste de hipótese para os 5 sinais analisados	30
Figura 3 – Exemplos de sinais industriais com diferentes tipos de distúrbios	33
Figura 4 – Comparativo da resposta em frequência de filtros passa-altas IIR (esquerda) e FIR (direita) de ordens 2 e 6.....	35
Figura 5 – Comparativo resposta em frequência de filtros IIR	36
Figura 6 – Especificação de filtros do tipo FIR.....	37
Figura 7 – Envelope espectral calculado sem filtragem e utilizando filtro de Butterworth de ordens 2 e 5.....	38
Figura 8 – Sinais com distúrbios e suas respectivas PSD	39
Figura 9 – Primeira etapa da segmentação de dados: classe inicialmente selecionada (em vermelho) e histograma (direita)	41
Figura 10 – Resultado da segmentação de dados (dados selecionados em vermelho)...	42
Figura 11 – Segmentação de dados aplicada a diferentes conjuntos de dados.....	44
Figura 12 – Comparativo entre a PSD calculada sem segmentação de dados e com a segmentação de dados	45
Figura 13 – Resultado da aplicação de remoção de tendências lineares (PSD à direita)	46
Figura 14 – Sinais no domínio do tempo com respectivos valores de desvio padrão e CV	49
Figura 15 – Dados que passaram por tratamento (segmentação de dados).....	50
Figura 16 – Desvio padrão e coeficiente de variação de sinais analisados	51
Figura 17 – Coeficiente de variação e o limiar de 0,7% aplicado aos sinais analisados	52
Figura 18 – Ilustração do algoritmo de seleção de intervalos	54
Figura 19 – Resultado da seleção de intervalos de frequência para diferentes valores de suavização.....	55
Figura 20 – Resultado da seleção automática de intervalos de frequência e da aplicação do limiar de energia para descartar intervalos	57
Figura 21 – PSD dos sinais analisados após utilizar a seleção automática de intervalos de frequência e o descarte de intervalos de frequência	58
Figura 22 – PSD de sinais sintéticos para diferentes valores de SNR.....	59
Figura 23 – Análise de energia e SNR	60
Figura 24 – Resultado da segmentação de dados (em vermelho dados descartados) e PSD após a segmentação de dados	62
Figura 25 – PSD resultante após a segmentação e a filtragem.....	63
Figura 26 – Envelope espectral calculado após tratamento (segmentação e filtragem), resultado da seleção automática de intervalos e resultado do descarte de intervalos através do limiar de energia	64
Figura 27 – CV dos sinais analisados e limiar (esquerda) e envelope espectral calculado com todos os sinais e apenas com os sinais selecionados pelo limiar do CV (direita) ..	65
Figura 28 – Comparativo entre a seleção automática e a manual (I)	67

Figura 29 – Comparativo entre a seleção automática e a manual (II)	67
Figura 30 - Comparativo entre a seleção automática e a manual (III)	68
Figura 31 – Dados originais sem tratamento (18h)	69
Figura 32 – Envelope espectral e intervalos de frequência selecionados (sombreado azul escuro) nas 5 análises para limiar de energia de 5% e 20%	71
Figura 33 – Resultado da detecção nas 5 análises para limiar de energia de 5% e de 20%	72
Figura 34 – Gráficos de repetibilidade dos resultados para o limiar de energia de 20%	73

Lista de Tabelas

Tabela 1 – Resultado relação entre amplitude dos picos do envelope calculado utilizando correlação cruzada e envelope somando espectros.....	26
Tabela 2 – Resultado da detecção de oscilações utilizando todos os sinais e apenas os sinais seleccionados pelo limiar do CV.....	66
Tabela 3 – Sinais seleccionados e descartados pelo limiar do CV nas 5 análises.....	70

Lista de Siglas

ACF	<i>Autocorrelation Function</i>
CV	Coeficiente de Variação
CWT	<i>Constinuous Wavelet Transform</i>
DCT	<i>Discrete Cossine Transform</i>
DTFT	<i>Discrete Time Fourier Transform</i>
DWT	<i>Discrete Wavelet Transform</i>
EMD	<i>Empirical Mode Decomposition</i>
FFT	<i>Fast Fourier Transform</i>
FIR	<i>Finite Impulse Response</i>
ICA	<i>Independent Component Analysis</i>
IIR	<i>Infinite Impulse Response</i>
ITD	<i>Intrinsic Time-Scale Decomposition</i>
OCI	<i>Oscillation Contribution Index</i>
PCA	<i>Principal Component Analysis</i>
PSD	<i>Power Spectral Density</i>
RZC	<i>Robust Zero Crossing Intervals Clustering</i>
SPCA	<i>Spectral Principal Component Analysis</i>

Sumário

1	Introdução.....	15
1.1	Objetivos	18
1.1.1	Objetivo Geral	18
1.1.2	Objetivos Específicos	18
1.2	Estrutura da Dissertação	18
2	O Método do Envelope Espectral.....	19
2.1	Estimação do Espectro	20
2.2	O Envelope Espectral	23
2.3	Efeito da correlação entre os sinais no Envelope Espectral	24
2.4	Teste de Hipótese	27
3	Contribuições para o método do Envelope Espectral.....	31
3.1	Pré – tratamento dos dados.....	31
3.1.1	Filtragem	33
3.1.2	Segmentação de dados	40
3.1.3	Remoção de tendências lineares.....	45
3.2	Impacto das oscilações nos sinais analisados.....	47
3.3	Seleção automática de intervalos de frequência.....	52
3.4	Proposta de utilização de um limiar de energia.....	55
3.4.1	Cálculo da energia.....	55
3.4.2	Relação sinal/ruído e energia	58
4	Estudos de caso	61
4.1	Pré-tratamento	61
4.2	Impacto das oscilações	64
4.3	Seleção automática de intervalos de frequência.....	66
4.4	Repetibilidade	68
5	Conclusão	75
	Referências Bibliográficas.....	76
6	APÊNDICE – Rotinas e códigos desenvolvidos em Matlab.....	79
6.1	Cálculo do CV e descarte de sinais com baixo impacto.....	80
6.2	Normalização, filtragem e segmentação de dados	81
6.3	Função para estimação do espectro e cálculo do envelope espectral	84
6.4	Função para seleção automática dos intervalos de frequência e descarte de intervalos utilizando limiar de energia.....	86

6.5	Função para realizar a indicação de sinais que possuem oscilações utilizando teste de hipótese	89
-----	--	----

1 Introdução

Oscilações são problemas comumente encontrados em diferentes tipos de processos industriais (JIANG; SHOUKAT CHOUDHURY; SHAH, 2007). Pesquisas sobre o monitoramento e desempenho de processos indicam que 60% dos controladores da indústria apresentam algum problema relacionado a desempenho (BIALKOWSKI, 1993; ENDER, 1993) e alguns desses resultam em oscilações. As causas de tais oscilações são diversas, desde agarramentos em válvulas de controle, controladores desajustados ou mal sintonizados e outros problemas relacionados a desgaste ou defeito em equipamentos e componentes do processo, além de distúrbios oscilatórios (KARRA; KARIM, 2009).

A detecção e diagnóstico de oscilações é de grande importância devido à frequência com que esses distúrbios estão presentes em sinais de processos industriais. A detecção de oscilações pode ser um grande auxílio para a equipe de operação de uma planta identificar algum problema que reduz o desempenho do processo e consequentemente da perda de qualidade do produto final. Sabe-se que esses distúrbios podem se propagar por toda a planta provocando aumento da variabilidade dos sinais (JIANG; SHOUKAT CHOUDHURY; SHAH, 2007; THORNHILL; COX; PAULONIS, 2003), e o impacto desses distúrbios no custo de produção é proporcional a amplitude do desvio em torno do ponto de operação (SHINSKEY, 1990).

Identificar defeitos em equipamentos pode auxiliar a equipe de manutenção a planejar a substituição dos mesmos e a executar a manutenção preventiva antes que ocorra um problema que cause parada do processo. Outro aspecto importante a ser lembrado está relacionado a segurança dos funcionários. A substituição preventiva de equipamentos defeituosos pode evitar danos mais graves e acidentes com implicações que afetem recursos materiais e humanos.

Nesse contexto, foram propostos diversos métodos para solucionar o problema de detecção de oscilações em processos industriais. Entre eles, pode ser citado o método Transformada Discreta do Cosseno (*Discrete Cossine Transform* – DCT) que tem como objetivo isolar as componentes de frequência de uma série temporal e verificar a regularidade dos cruzamentos por zero das componentes independentes resultantes (LI et al., 2010). O método foi proposto para séries temporais univariadas e aplicado a dados simulados e dados industriais. Posteriormente, o método proposto foi melhorado por Wang et al., (2013). Foram feitos testes através de estudos de caso de dados industriais e,

mesmo com as melhorias feitas, o autor ressaltou que surgiram dificuldades e problemas na implementação online.

O método de detecção baseado na Função de Autocorrelação (*Autocorrelation Function* – ACF) foi proposto por Thornhill et al. (2003) utilizando conceitos abordados em Seborg (1999) sobre a utilização de ACF. No método baseado em ACF, verifica-se os cruzamentos por zero da função de autocorrelação do sinal, já que esta função evidencia a característica periódica do sinal e atenua ruídos. Verificam-se então a regularidade do sinal, a taxa de decaimento e a força do sinal. Posteriormente o método foi melhorado por Karra e Karim (2009) com a adição do cálculo da Densidade Espectral de Potência (*Power Spectral Density* – PSD) na metodologia com o intuito de identificar frequências dominantes e filtrar as faixas de frequência de interesse.

A Análise das Componentes Principais (*Principal Component Analysis* - PCA) é uma ferramenta matemática utilizada para realizar uma redução na dimensionalidade em um conjunto de dados. Ela possui aplicações variadas em diversas áreas de estudo, como por exemplo em detecção de oscilações, proposta por Thornhill et al. (2002). No trabalho em questão, o PCA foi aplicado ao espectro de sinais, chamado de Análise Espectral das Componentes Principais (*Spectral Principal Component Analysis* – SPCA). Uma outra abordagem proposta em Xia e Howell (2005) e Xia et al. (2005) usa o método Análise das Componentes Independentes (*Spectral Independent Component Analysis* – *Spectral ICA*) aplicado ao espectro de potência com o objetivo de encontrar uma representação linear dos dados na qual as componentes sejam estatisticamente independentes.

A decomposição de modo empírico (*Empirical Mode Decomposition* – EMD) foi proposta por Srinivasan et al.; (2007) em uma versão modificada para caracterização de oscilações. O objetivo da metodologia é identificar todos os cruzamentos por zero presentes nos dados utilizando EMD que decompõe os dados em uma porção chamada de função de modo intrínseco e uma componente contendo a tendência do segmento de dados com a vantagem de conseguir lidar com dados não estacionários. Posteriormente foram propostas melhorias na metodologia selecionando apenas as funções de modo intrínseco significativas e incluindo análises das funções de auto correlação e transformada de Fourier (SRINIVASAN; RENGASWAMY, 2012).

Uma abordagem online para solucionar o problema da presença de oscilações intermitentes foi proposta utilizando Decomposição Intrínseca em Escala de Tempo (*Intrinsic Time-Scale Decomposition* - ITD) em conjunto com Agrupamento Robusto dos Intervalos de Cruzamentos por Zero (*Robust Zero Crossing Intervals Clustering* - RZC)

(XIE et al., 2016). O método consiste na utilização de ITD para decompor o sinal em diferentes componentes de frequência e agrupamento via algoritmo *k-means* para eliminar cruzamentos por zero ruidosos.

Uma outra metodologia utilizada para a detecção de múltiplas oscilações em processos industriais é a aplicação de transformadas *wavelet*, que foi utilizada em Matsuo et al., (2003) e Guo et al., (2014). As transformadas *wavelet* tem o objetivo de decompor o sinal em diferentes componentes de frequência. A Transformada Discreta de *Wavelet* (*Discrete Wavelet Transform – DWT*) e a Transformada Contínua de *Wavelet* (*Continuous Wavelet Transform – CWT*) são utilizadas em Guo et al., (2014) em conjunto com ACF com o objetivo obter um método automático de detecção de oscilações.

O envelope espectral proposto pela primeira vez em Stoffer et al. (1993), melhorado posteriormente em McDougall et al. (1997), tem como objetivo obter uma transformação linear ótima de séries temporais que evidencie a natureza periódica dos dados e atenuar os ruídos. Pode-se obter então uma única curva que representa a maior porção de potência em cada frequência, chamada de envelope, facilitando a visualização de todas as oscilações presentes nos dados. Em Jiang et al. (2007) a metodologia foi aplicada a dados industriais, utilizando-se também o teste de hipótese para categorizar oscilações presentes em vários sinais e um índice proposto pelos autores para analisar a causa raiz das oscilações. A utilização desta metodologia requer que os sinais utilizados tenham sido amostrados simultaneamente.

Um trabalho foi publicado em (FARIA et al., 2017) com o objetivo de realizar a detecção de oscilações através do envelope ponderado, obtido através da soma das PSD ponderadas pelos seus respectivos CV. Também foi utilizada a seleção automática de intervalos de frequência em conjunto com o descarte de intervalos utilizando limiar de energia. A indicação de sinais em que as oscilações estavam presentes foi realizada através da verificação da regularidade dos cruzamentos por zero das funções de autocorrelação. O trabalho de Faria et al., (2017) serviu como base para o desenvolvimento das melhorias propostas nesta dissertação.

A motivação deste trabalho é apresentar contribuições para a metodologia de detecção de oscilações em processos industriais baseada no envelope espectral. Foram implementadas análises adicionais em relação a metodologia apresentada na literatura, são elas: cuidados no tratamento dos dados industriais contaminados por distúrbios e ruídos, utilização do coeficiente de variação para descartar sinais de baixo impacto na análise, seleção automática de intervalos de frequência, utilização de limiar de energia

para descartar intervalos de frequência relacionados a oscilações irregulares e indicação de sinais contendo oscilações relevantes através do teste de hipótese.

1.1 Objetivos

1.1.1 Objetivo Geral

O objetivo geral desta dissertação é propor melhorias ao método de detecção de oscilações baseado no envelope espectral.

1.1.2 Objetivos Específicos

- Analisar o pré-tratamento para aplicar o método do envelope espectral;
- Propor limiares para melhorar a repetibilidade dos resultados e eliminar oscilações de baixo impacto;
- Automatizar a escolha das faixas de frequência do envelope que devem ser investigadas.

1.2 Estrutura da Dissertação

O presente trabalho se divide da seguinte maneira: o Capítulo 2 apresenta um histórico mostrando a evolução da metodologia do envelope espectral desde a primeira vez que foi proposto, e a teoria relacionada a estimação do espectro e ao envelope espectral. Nesse capítulo também é feita uma análise sobre o efeito da correlação entre sinais no envelope espectral e, por último, uma descrição da utilização de um teste de hipótese para indicar os sinais onde existem oscilações. O Capítulo 3 apresenta as contribuições do presente trabalho para a metodologia do envelope espectral: os tratamentos adequados para os diferentes distúrbios encontrados nos dados, proposta de utilização de uma medida de impacto das oscilações para descartar sinais, proposta para seleção automática de intervalos de frequência e utilização de um limiar de energia para descartar oscilações irregulares. O Capítulo 4 apresenta os estudos de caso utilizados para ilustrar a metodologia proposta e avaliar o desempenho da mesma ao detectar oscilações em dados industriais com diferentes características. Também foi verificada a repetibilidade dos resultados em análises subsequentes. Por último, o Capítulo 5 contém as conclusões obtidas a partir das contribuições propostas na dissertação e discussão dos resultados.

2 O Método do Envelope Espectral

O objetivo deste capítulo é descrever a teoria relacionada a metodologia do envelope espectral. Na primeira seção serão apresentadas as equações e teoria utilizada para estimação do espectro. Na segunda seção são detalhadas as equações necessárias para calcular o envelope espectral. Na terceira seção são expostas análises com o objetivo de investigar o efeito da correlação cruzada dos sinais no envelope espectral. Na última seção do capítulo é detalhado o teste hipótese aplicado ao envelope espectral com o objetivo de indicar em quais sinais as oscilações identificadas no envelope estão presentes.

O método do envelope espectral utilizado para analisar a natureza periódica de dados reais foi apresentado pela primeira vez em Stoffer et al. (1993), e surgiu da escassez de metodologias para investigar características periódicas de séries temporais no domínio da frequência. No artigo em questão, foram atribuídos valores para cada categoria antes de se realizar a análise do espectro, baseando-se no fato de que o envelope pode ser aplicado na série temporal original e também na série temporal transformada. Porém, esses valores foram atribuídos de forma arbitrária e verificou-se que o espectro resultante depende deles. A metodologia proposta foi aplicada a dados de biologia molecular.

Em McDougall (1997) melhorou-se o envelope proposto por Stoffer et al. (1993) e o conceito de escalonamento ótimo para o cálculo do envelope foi explorado com o propósito de evidenciar a natureza periódica dos dados e atenuar ruídos, substituindo assim a atribuição arbitrária de valores. O trabalho em questão também estendeu as análises a séries temporais multivariadas, que até então vinham sendo utilizadas apenas em séries com uma variável. Os resultados foram demonstrados em dados de Produto Interno Bruto (PIB), valores de retorno de investimentos de empresas, entre outros.

Em Stoffer (1999) o objetivo principal foi identificar oscilações comuns em diversos sinais para analisar séries temporais multivariadas. As análises foram demonstradas utilizando dados da área de saúde. Toda a metodologia do envelope já desenvolvida nos artigos anteriores é apresentada em Wendt et al. (2000). As aplicações do envelope espectral foram testadas em séries temporais e foram fornecidos exemplos utilizando dados reais.

Em Jiang et al (2007), todos os conceitos já explorados nos trabalhos anteriores foram aplicados a processos industriais com o objetivo de detectar oscilações. Os resultados

foram comparados com a técnica de Análise de Componentes Principais aplicada ao espectro (*Spectral Principal Component Analysis* - SPCA). Foi adicionada também a análise da provável causa raiz da oscilação através de um Índice de Contribuição de Oscilações (*Oscillation Contribution Index* - OCI) proposto pelos autores.

2.1 Estimação do Espectro

A primeira etapa necessária para o cálculo do envelope espectral é a transformação dos dados do domínio do tempo para o domínio da frequência, através do cálculo do espectro. Na prática, os dados têm natureza aleatória pois não se pode determinar através do histórico dos dados qual vai ser o comportamento futuro do mesmo. Em geral, dados aleatórios possuem energia média finita e a representação de sua potência na frequência é chamada de Densidade Espectral de Potência (*Power Spectral Density* - PSD). No entanto, dados aleatórios amostrados (sequências discretas) não possuem energia finita, portanto não é possível calcular a Transformada de Fourier de Tempo Discreto (*Discrete Time Fourier Transform* – DTFT) desses sinais.

Além disso, os dados têm disponibilidade limitada, ou seja, uma janela de dados é obtida através da coleta de amostras da quantidade total dos dados existente. Por essa razão a PSD de um sinal deve ser obtida através de uma estimativa do espectro a partir dos dados disponíveis.

Existem diversas formas de se obter uma estimação do espectro. As metodologias de estimação do espectro podem ser divididas em duas categorias: paramétricas e não-paramétricas. A estimação paramétrica consiste na obtenção de um modelo espectral baseado nos dados obtidos, portanto, trata-se de um problema de estimação dos parâmetros de um modelo. Já a estimação não-paramétrica (clássica) é uma abordagem diferente em que se utiliza uma janela móvel de dados para estimar qual a potência do sinal em cada frequência (KOOPMANS, 1995; STOICA, PETRE; MOSES, 2005).

Entre as metodologias de estimação paramétrica as mais conhecidas e simples são o Periodograma e o Correlograma. Em certas circunstâncias as duas metodologias são equivalentes. A vantagem dessas metodologias é o fornecimento de estimativas simples de alta resolução, porém as estimativas não são consistentes já que possuem alta variância e a variância não diminui à medida que se aumenta a quantidade de dados utilizada. Versões modificadas do Periodograma e do Correlograma, que fazem uso de filtragem ou

suavização, foram desenvolvidas com o objetivo de corrigir o problema de alta variância, porém, possuem a desvantagem de uma resolução reduzida.

No presente trabalho as funções de autocorrelação e correlação cruzada foram utilizadas para obter a matriz Hermitiana P_x (necessária para obtenção do envelope) que contém a estimação do espectro de cada sinal (diagonal principal da matriz) e o espectro cruzado entre os sinais (outros elementos da matriz). A utilização das funções de correlação (Correlograma) ao invés do uso do sinal de forma direta (Periodograma) para estimar o espectro se justifica pela característica de redução de ruído e evidência das características oscilatórias do sinal.

Considerando um conjunto de dados amostrados no domínio do tempo tal que a matriz X de dados apresentada na Equação (1) seja composta de m sinais $x(t)$, cada sinal com n amostras.

$$X = \begin{bmatrix} x_1(t) & \dots & x_1(t+n) \\ \vdots & \ddots & \vdots \\ x_m(t) & \dots & x_m(t+n) \end{bmatrix} \quad (1)$$

Sejam esses dados normalizados, então a correlação cruzada entre dois sinais pode ser definida através da Equação (2), onde x_1 e x_2 são os dois sinais em que se pretende obter a correlação, o operador E é o valor esperado e o operador $(\cdot)^*$ representa o complexo conjugado.

$$r_{x_1x_2}(k) = E\{x_1(t)x_2^*(t-k)\} \quad (2)$$

De maneira análoga, a autocorrelação de um sinal pode ser calculada utilizando a Equação (2) e substituindo o sinal x_2 pelo próprio sinal x_1 .

A Equação (2) prevê requisitos para que seja válida a utilização da mesma na obtenção da correlação entre sinais ou da autocorrelação. Vale ressaltar que, na prática, a disponibilidade de dados é limitada, portanto tem-se disponível apenas intervalos de dados de uma série completa. Além disso, para a estimação da correlação os dados devem ser normalizados. Os dados aleatórios sendo utilizados também devem ser estacionários. Dados podem ser considerados estacionários se sua média e variância não variam sistematicamente no tempo (GUJARATI; PORTER, 2011). Tendências e mudanças de

ponto de operação, por exemplo, fazem com que os dados não possam ser considerados estacionários.

Aplicando a Transformada Rápida de Fourier (*Fast Fourier Transform* – FFT) na função de correlação cruzada apresentada na Equação (2), obtém-se a estimação da função densidade de potência do espectro cruzado (STOICA, PETRE; MOSES, 2005), conforme a Equação (3).

$$\phi_{x_1x_2}(\omega) = \frac{1}{n} \sum_{k=1}^n r_{x_1x_2}(k) e^{-i\omega k} \quad (3)$$

De forma similar, aplicando a FFT na função de autocorrelação de um único sinal, estima-se a densidade espectral de potência do sinal através Equação (4).

$$\phi_{x_1}(\omega) = \frac{1}{n} \sum_{k=1}^n r_{x_1}(k) e^{-i\omega k} \quad (4)$$

Estimando-se o espectro de cada sinal e dos espectros cruzados entre os sinais, pode-se obter a matriz Hermitiana, $\Phi(\omega)$, apresentada em (5). Os elementos da diagonal principal da matriz contêm os espectros estimados dos m sinais e os elementos restantes da matriz são os espectros cruzados entre os sinais.

$$\Phi(\omega) = \begin{bmatrix} \phi_{x_1}(\omega) & \dots & \phi_{x_1x_m}(\omega) \\ \vdots & \ddots & \vdots \\ \phi_{x_mx_1}(\omega) & \dots & \phi_{x_m}(\omega) \end{bmatrix} \quad (5)$$

A utilização da FFT do sinal autocorrelacionado como forma de se obter uma estimativa para o espectro possui a desvantagem de ter uma variância alta e que não diminui à medida que se aumenta a quantidade de amostras n , ou seja, a estimativa não é consistente. Esse problema pode ser resolvido com a utilização de uma técnica de suavização. Neste trabalho a metodologia escolhida foi uma solução simples e básica e a mesma adotada em Jiang et al. (2007), o filtro de média móvel. A técnica consiste em obter a média móvel da estimativa $\Phi(\omega)$ através da Equação (6), onde h_j são os pesos simétricos obtidos através de: $h_j = \frac{r-|j|+1}{(r+1)^2}$, $|j| = 1, 2, \dots, r$, sendo r o fator de suavização.

$$\hat{P}_x(\omega) = \sum_{j=-r}^r h_j \Phi(\omega) \quad (6)$$

De posse da matriz \hat{P}_x , que contém o espectro estimado, o envelope espectral pode ser estimado, como será descrito na próxima seção.

2.2 O Envelope Espectral

Seja a matriz de dados da definida em (1), sendo V_x a matriz de covariância obtida a partir da matriz X e $P_x(\omega)$ a matriz contendo o espectro estimado. Seja também a série temporal escalonada $g(t, \beta) = \beta^* x(t)$ resultante da combinação linear das linhas de X . O vetor coluna β , de tamanho m , representa o vetor de escalonamento da série temporal. A covariância desta série temporal transformada pode ser definida como $V_g(\beta) = \beta^* V_x \beta$ e a densidade espectral de potência da série temporal escalonada é definida por $P_g(\omega, \beta) = \beta^* P_x(\omega) \beta$. O conceito do envelope espectral pode então ser definido a partir da Expressão (7).

$$\lambda(\omega) \triangleq \sup_{\beta \neq 0} \left\{ \frac{P_g(\omega, \beta)}{V_g(\beta)} \right\} = \sup_{\beta \neq 0} \left\{ \frac{\beta^* P_x(\omega) \beta}{\beta^* V_x \beta} \right\} \quad (7)$$

O operador \sup (supremo), representa o maior valor do resultado das operações matemáticas contidas dentro das chaves. O envelope espectral representa a maior porção de potência ou variância contida numa determinada frequência para a série escalonada pelo vetor de escalonamento β . O vetor de escalonamento β que resulta na obtenção $\lambda(\omega)$ é chamado de vetor de escalonamento ótimo β . O vetor de escalonamento é tal que $\beta^* V_x \beta = 1$, ou seja, a série temporal escalonada tem variância unitária. Como consequência, a Equação (7) pode ser reescrita no formato $\lambda(\omega) V_x \beta(\omega) = P_x(\omega) \beta(\omega)$ e o envelope espectral $\lambda(\omega)$ pode ser obtido através do maior autovalor associado à equação $|P_x(\omega) - \lambda(\omega) V_x| = 0$ e $\beta(\omega)$ os correspondentes autovetores.

Após o conceito do envelope espectral ter sido introduzido, sabe-se que, na prática o espectro deve ser estimado a partir dos dados e, de forma análoga, o envelope espectral também deve ser estimado a partir do espectro estimado. Portanto, a partir das propriedades e equação do envelope espectral (7), e de posse do espectro estimado $P_x(\omega)$

(Equação (6)), a covariância dos dados $\hat{V}_x = Cov(X)$ e $\hat{V} = Diag(\hat{V}_x)$, a estimação do envelope espectral pode ser obtida através da Expressão (8).

$$\hat{\lambda}(\omega) = \sup_{\beta \neq 0} \left\{ \frac{\hat{\beta}^* \hat{P}_x(\omega) \hat{\beta}}{\hat{\beta}^* \hat{V} \hat{\beta}} \right\} \quad (8)$$

Sabe-se que os dados utilizados são normalizados, portanto, a diagonal da covariância dos dados \hat{V} é unitária. E, como consequência disso, a restrição $\hat{\beta}^* \hat{V} \hat{\beta} = 1$ e $\hat{\beta}^* \hat{\beta} = 1$. Substituindo essas duas expressões na Equação (8) conclui-se que a estimação do envelope espectral pode ser simplesmente obtida através do maior autovalor de $\hat{P}_x(\omega)$ e o correspondente autovetor $\hat{\beta}$.

O envelope espectral possui alguns requisitos em relação aos dados que estão sendo utilizados. Assim como a maioria das metodologias de análise espectral, um requisito para se utilizar a metodologia do envelope espectral é que a série temporal utilizada deve ser estacionária (WENDT; TYLER; STOFFER, 2000). Segundo Gujarati e Porter (2011) uma série temporal pode ser considerada estacionária se sua média e variância não variam sistematicamente no tempo. Comportamentos não estacionário resultam em distúrbios. Além disso os sinais devem ser amostrados simultaneamente. Um outro aspecto relacionado aos requisitos da metodologia do envelope é a normalização dos dados. Em Jiang et al. (2007) são ressaltadas algumas vantagens de se utilizar dados normalizados como: a simplificação dos cálculos e da interpretação do envelope espectral, a possibilidade de se comparar os envelopes espectrais de diferentes conjuntos de dados, a dimensão dos elementos do vetor $\beta(\omega)$ torna-se também mais significativa e fácil de interpretar, e uma estimação precisa da correlação entre dados demanda que os mesmos sejam normalizados. Outra vantagem de se utilizar dados normalizados está no cálculo do próprio envelope espectral, que passa a ser obtido diretamente a partir do maior autovalor da matriz do espectro estimado em cada frequência.

2.3 Efeito da correlação entre os sinais no Envelope Espectral

Nas seções anteriores foram desenvolvidas as equações e a metodologia para se estimar o espectro e o envelope espectral. Existem diversas metodologias para se estimar a PSD de um sinal, porém a metodologia escolhida neste trabalho usa a função de autocorrelação e correlação cruzada, com vantagens tais como a redução do efeito do ruído.

Nas referências do envelope espectral utilizadas como ponto de partida para o desenvolvimento deste trabalho não foram investigados os efeitos de diferentes metodologias para se estimar o espectro e o envelope espectral o que levanta uma pergunta: qual o efeito da correlação entre os sinais no envelope espectral e como o espectro cruzado influencia o envelope? Neste trabalho as funções de correlação são utilizadas não só para gerar o espectro de cada sinal, mas também para calcular o espectro cruzado entre os sinais, ambos contidos na matriz \hat{P}_x .

Para investigar esse efeito uma análise pode ser feita comparando-se duas formas de obtenção do envelope espectral. Uma delas é apresentada nas referências já citadas neste trabalho (JIANG; SHOUKAT CHOUDHURY; SHAH, 2007; WENDT; TYLER; STOFFER, 2000), na qual o envelope é obtido através maior autovalor da matriz do espectro estimado, \hat{P}_x , o que significa obter a maior porção de potência em cada frequência. Como a matriz \hat{P}_x contém o espectro cruzado dos sinais nos elementos fora de sua diagonal principal, sabe-se que haverá influência da correlação dos sinais no envelope espectral.

Uma outra possível abordagem para investigar o efeito da correlação é calcular o envelope sem o efeito da correlação dos sinais, sem o espectro cruzado, que pode ser obtido a partir da soma dos elementos da diagonal principal da matriz \hat{P}_x , ou seja, a estimativa da PSD de cada sinal autocorrelacionado. Uma metodologia similar proposta em Faria et al., (2017), onde o envelope foi obtido através da soma da PSD de cada sinal ponderada pelo CV. Dessa forma o efeito da correlação cruzada entre os sinais pode ser verificado através da comparação dos dois envelopes espectrais.

Na análise a seguir comparou-se o envelope espectral com o envelope obtido através da soma da PSD de dois sinais, aumentando-se a correlação entre eles. Para isso foram gerados dois sinais oscilatórios $y_1(t) = \text{sen}(\omega_1 t)$ e $y_2(t) = \text{sen}(\omega_2 t) + ay_1$. O coeficiente a foi variado de 0,1 a 1 para avaliar o efeito da correlação entre os dois sinais.

Nota-se através da Tabela 1 e da Figura 1 que, à medida que a correlação entre os sinais aumenta, o efeito da mesma não aparece no envelope utilizando a soma das PSD para $a = 0,1$ e $a = 0,35$. Com o valor de $a = 0,6$ existe uma grande participação da frequência ω_1 no sinal y_2 , portanto, devido a correlação alta entre y_1 e y_2 espera-se um grande efeito no envelope, que pode ser notado no envelope utilizando a correlação cruzada, mas não no envelope utilizando a soma dos espectros. Na Tabela 1 a terceira

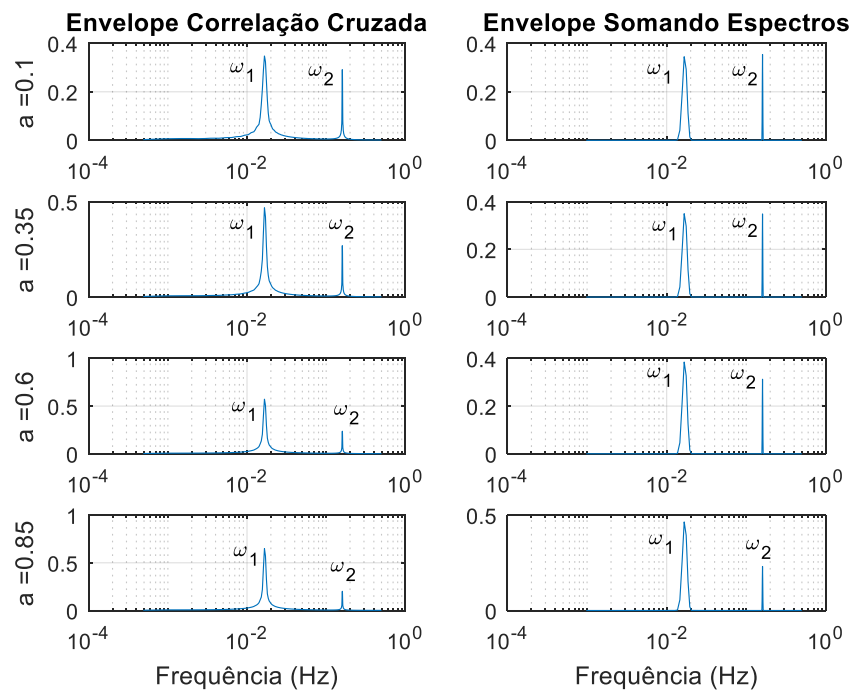
coluna representa a relação entre as amplitudes dos picos utilizando a correlação ($\phi(\omega_1)/\phi(\omega_2)$) e a quarta coluna representa a relação entre amplitudes dos picos quando o envelope é calculado através da soma ($\lambda_s(\omega_1)/\lambda_s(\omega_2)$).

Tabela 1 – Resultado relação entre amplitude dos picos do envelope calculado utilizando correlação cruzada e envelope somando espectros

a	Correlação Cruzada	$\phi(\omega_1)/\phi(\omega_2)$	$\lambda_s(\omega_1)/\lambda_s(\omega_2)$
0,1	0,099	1,103	0,971
0,35	0,33	1,704	1
0,60	0,515	2,478	1,226
0,85	0,648	3,201	2,001

Fonte: O próprio autor.

Figura 1 – Envelope calculado utilizando correlação cruzada e envelope somando espectros



Fonte: O próprio autor.

Com os resultados expostos, conclui-se que, numa situação em que não existe correlação entre os sinais ou a correlação é pequena (Figura 1, para $a=0,1$), calcular o envelope somando os espectros ou através da correlação cruzada produz resultados semelhantes. Porém, quando existe correlação entre os sinais, é mais representativo utilizar apenas o envelope espectral obtido através da correlação cruzada entre os sinais.

2.4 Teste de Hipótese

O cálculo do envelope espectral fornece uma única curva que representa o espectro a partir de um conjunto de dados com múltiplos sinais. O envelope desempenha então uma importante função visual da distribuição da potência no espectro. Todavia, apenas a curva do envelope não fornece indicação de quais sinais contêm as oscilações representadas pelos picos. Essa indicação dos sinais e oscilações faz parte da metodologia proposta inicialmente em Wendt et al. (2000), trabalho no qual a utilização do vetor de escalonamento ótimo para indicação de sinais foi apresentada.

Em Jiang et al. (2007) foi desenvolvido o teste de hipótese utilizando o vetor de escalonamento ótimo para indicação dos sinais onde existem oscilações. No presente trabalho foi adotada a mesma metodologia citada anteriormente, porém a verificação do resultado do teste de hipótese foi feita de forma computacional e automática, e não manualmente como em Jiang et al. (2007).

A utilização do teste de hipótese para verificar a existência de oscilações nos sinais parte do vetor de escalonamento ótimo $\hat{\beta}(\omega)$. Os elementos de $\hat{\beta}(\omega)$ fornecem a medida da contribuição de cada série temporal ou sinal para o envelope. Uma vez identificado o pico no envelope espectral, em uma determinada frequência, pode-se investigar o valor do escalonamento ótimo na frequência. O maior valor de escalonamento ótimo dentre todos os m sinais para esta frequência indica uma maior contribuição daquele sinal no envelope. Porém, verificar se determinada oscilação existe ou não naquele sinal implica em realizar o teste estatístico nos valores de $\hat{\beta}(\omega)$ para verificar se os escalonamentos correspondentes são estatisticamente significativos.

Antes de aplicar o teste de hipótese em $\hat{\beta}(\omega)$, deve ser abordado o motivo e significado da aplicação do teste estatístico nesse caso. Sabe-se que, se o espectro calculado $\hat{P}_x(\omega)$ for uma estimativa consistente, e, se $\hat{\lambda}(\omega)$ for um autovalor distinto pode-se concluir que a expressão $v_n[\hat{\beta}(\omega) - \beta(\omega)]$ converge para uma distribuição normal multivariada complexa quando $n \rightarrow \infty$, onde $v_n = (\sum_{j=-r}^r h_j^2)^{\frac{1}{2}}$ é derivado da Equação (6), em que se obtém o espectro estimado após a suavização utilizando a média móvel simétrica (JIANG; SHOUKAT CHOUDHURY; SHAH, 2007). Antes de se definir a expressão do teste de hipótese, deve ser calculada a covariância assintótica de $\hat{\beta}(\omega)$. Para isso, serão utilizados não apenas o maior autovalor $\hat{\lambda}(\omega)$ e respectivo maior

autovetor $\hat{\beta}(\omega)$ para cada frequência, mas também todos os autovalores $\lambda_z(\omega)$ e todos respectivos autovetores $\beta_z(\omega)$ que possuem dimensão máxima m . $\lambda_z(\omega)$ e $\beta_z(\omega)$ são ordenados em ordem decrescente, portanto os primeiros elementos de $\lambda_z(\omega)$ e $\beta_z(\omega)$ são respectivamente $\hat{\lambda}(\omega)$ e $\hat{\beta}(\omega)$.

A matriz de covariância assintótica $V_\beta(\omega)$ pode ser obtida através da Equação (9), sendo $V_1(\omega)$ a primeira parte da equação utilizada para calcular a matriz de covariância assintótica (Equação (10)) e $V_2(\omega)$, a segunda parte, obtido através da Equação (11). A variável v_n já foi definida no parágrafo anterior e $\hat{\lambda}(\omega)$ contém apenas o maior autovalor para cada frequência ω .

$$V_\beta(\omega) = V_1(\omega)V_2(\omega) \quad (9)$$

$$V_1(\omega) = v_n^{-2} * \hat{\lambda}(\omega) \quad (10)$$

$$V_2(\omega) = \sum_{z=2}^m \lambda_z(\omega) [\hat{\lambda}(\omega) - \lambda_z(\omega)]^{-2} \beta_z \beta_z^* \quad (11)$$

A partir da Equação (9) pode ser obtido $\sigma_j(\omega)$ calculando-se a diagonal da matriz $V_\beta(\omega)$ para cada frequência. De posse de todos os elementos para realizar o teste de hipótese, a estatística pode ser definida através da Equação (12).

$$T(\omega) = \frac{2|\hat{\beta}(\omega)|^2}{\sigma_j(\omega)} \quad (12)$$

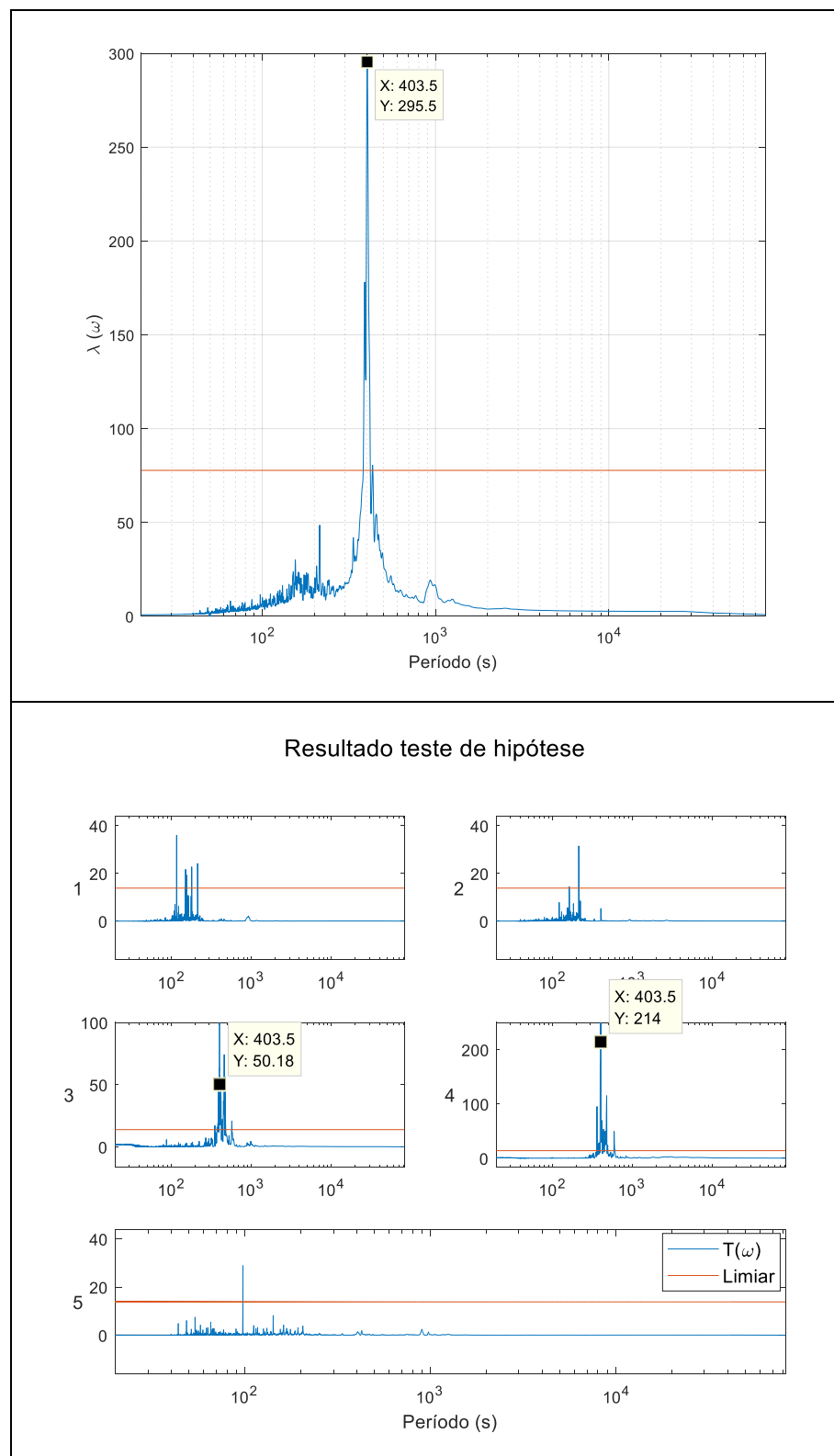
O teste pode ser feito para cada frequência ω de modo que, o resultado $T(\omega)$ deve ser comparado a distribuição qui-quadrado e com 2 graus de liberdade com o nível de confiança de 99,9% obtendo-se um limiar de aproximadamente 13,82. O motivo para se utilizar 2 graus de liberdade pode ser explicado pelo número de variáveis escolhidas livremente após terem sido impostas as restrições.

O resultado do teste pode ser avaliado da seguinte forma: caso $T(\omega) > \chi_2^2(\alpha)$, a hipótese nula de que $\hat{\beta}(\omega) = 0$ pode ser rejeitada, o que significa que existe oscilação na frequência sendo analisada. Se $T(\omega) < \chi_2^2(\alpha)$, a hipótese nula de que $\hat{\beta}(\omega) = 0$ pode ser aceita, o que significa que não existe oscilação na frequência sendo analisada. O

procedimento pode ser automatizado verificando se o resultado do teste de hipótese ultrapassa o limiar de 13,82 para $\omega = \omega_1$, sendo ω_1 pertencente ao intervalo de frequência, indicando que existe uma oscilação naquele sinal. Cada intervalo de frequência do envelope que indica uma oscilação é verificado realizando uma varredura em $\hat{\beta}(\omega)$ em cada sinal e verificando o resultado do teste de hipótese para uma amostra dentro do intervalo de frequência selecionado, posteriormente seguindo para outros intervalos de frequência.

A Figura 2 mostra um exemplo usando 5 sinais industriais em que o teste de hipótese foi realizado. Na curva superior o envelope espectral indica uma oscilação presente em 403s. No gráfico inferior são mostradas as curvas que expressam o resultado do teste de hipótese $T(\omega)$ realizado em cada sinal (Equação (12)), na cor azul, e, em vermelho, o limiar $\chi^2_2(\alpha)$ com 2 graus de liberdade e nível de confiança de 99,9%. Pode-se verificar que os sinais 3 e 4 passaram no teste de hipótese na frequência analisada, indicando a presença da oscilação com período 403s nos sinais 3 e 4.

Figura 2 – Envelope espectral e correspondente resultado do teste de hipótese para os 5 sinais analisados



Fonte: O próprio autor

3 Contribuições para o método do Envelope Espectral

Após ter sido apresentada a teoria relacionada ao método do envelope espectral, o objetivo do Capítulo 3 é apresentar as contribuições propostas por este trabalho tendo como base o método do envelope espectral.

Primeiramente será feita uma discussão sobre os tratamentos que devem ser aplicados a dados reais com o objetivo de viabilizar a aplicação da metodologia do envelope espectral. Serão mostrados exemplos de dados reais e dos diferentes tipos de distúrbios presentes nos dados, bem como as possíveis abordagens para tratar os diferentes tipos de distúrbios. Na segunda seção do capítulo o objetivo é propor a utilização de uma medida de impacto das oscilações para descartar sinais das análises. Na terceira seção será apresentada uma proposta para selecionar automaticamente intervalos de frequência do envelope espectral para substituir a seleção manual dos picos. Na última seção será proposta a utilização de um limiar baseado na energia da oscilação com o propósito de descartar intervalos de frequência de oscilações irregulares.

Com as contribuições propostas nesse capítulo, foi desenvolvido um algoritmo de detecção de oscilações baseado na metodologia do envelope espectral que segue os passos abaixo:

1. Pré – tratamento dos dados;
2. Cálculo do impacto das oscilações e descarte de sinais com baixo impacto;
3. Cálculo do espectro;
4. Cálculo do Envelope Espectral;
5. Seleção automática de intervalos de frequência do envelope;
6. Cálculo da energia dos intervalos e descarte de intervalos de frequência com baixa energia;
7. Indicação de quais sinais possuem oscilações através de teste de hipótese realizado para uma amostra de ω pertencente ao intervalo de frequência selecionado.

3.1 Pré – tratamento dos dados

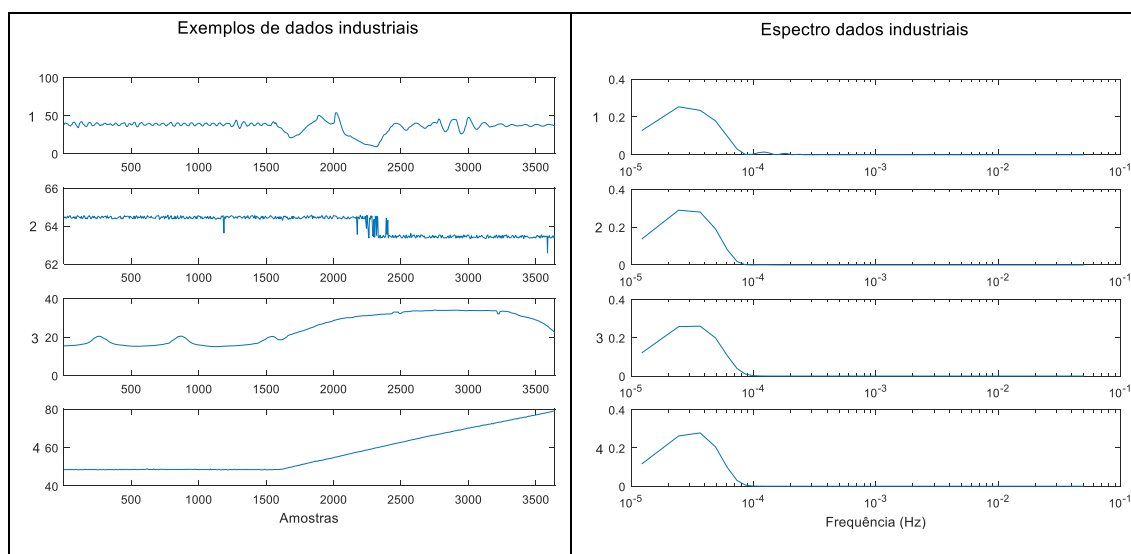
No Capítulo 2 foram listados alguns requisitos para utilização do método do envelope espectral. Entre eles a necessidade de a série temporal ser estacionária e os vetores de

dados dos sinais terem instantes de tempo correspondentes, ou seja, amostras de diferentes sinais devem corresponder ao mesmo instante de tempo.

Dados industriais em geral possuem ruídos de medição, diversos tipos de distúrbio, tendências, *outliers* e mudanças de ponto de operação. A presença dessas variações nos sinais pode fazer com que eles não sejam mais considerados sinais estacionários, causando mudanças na média e no desvio padrão do sinal. Por esse motivo torna-se necessária a utilização de pré – tratamento nos dados. O objetivo desta seção é apresentar a variada gama de comportamentos e características presentes nos dados reais (industriais) que prejudicam ou inviabilizam a utilização do método do envelope espectral para detecção de oscilações, bem como propor tratamentos adequados para cada tipo de problema, de modo que os dados estejam dentro dos requisitos.

A Figura 3 mostra alguns exemplos de dados industriais obtidos a partir de sensores de diferentes processos em uma unidade estacionária de produção de óleo e gás. Nota-se, por exemplo, no sinal 1 um forte distúrbio. O sinal 2 é um exemplo de sinal que contém mudança de ponto de operação e também alguns *outliers*. O sinal 3 é oscilatório irregular e contém um distúrbio oscilatório de baixa frequência (apenas um semiperíodo). O sinal 4 contém um trecho estacionário seguido de uma rampa, ou seja, uma tendência linear. A Figura 3 contém algumas das alterações mais comuns presentes nos dados que estão sendo utilizados nesse trabalho, alterações que tem grande influência ao se obter o espectro dos sinais e que precisam ser tratadas. Como pode ser visto no espectro de cada sinal, as alterações em sua maioria produzem componentes de grande amplitude em baixa frequência no espectro. No sinal 1, onde existe claramente uma oscilação regular, esta oscilação acaba sendo ofuscada no espectro.

Figura 3 – Exemplos de sinais industriais com diferentes tipos de distúrbios



Fonte: O próprio autor

3.1.1 Filtragem

Existem situações em que filtros são necessários para eliminar ruídos ou distorções presentes nos sinais. Entre os muitos aspectos da detecção de oscilações abordados em Thornhill et al. (2003), um deles trata da automatização do processo de filtragem dos dados. No trabalho citado os autores propõem um algoritmo com o objetivo de redução do intervalo de frequência dos sinais investigados, baseado nas oscilações que já foram detectadas. O método consiste nos seguintes passos:

1. Pré-processamento dos dados descartando tendências não estacionárias com menos que 8 ciclos completos;
2. Aplicação do método de detecção de oscilações nos dados processados. Seleção dos períodos de qualquer oscilação detectada;
3. Ajuste das bordas do filtro utilizando as bordas do filtro de pré-processamento e as oscilações detectadas no passo anterior;
4. Repetir a análise para os novos intervalos de frequência, toda vez que for redefinido. Repetir o passo 3 para o caso de múltiplas frequências num mesmo sinal. Parar quando não houver casos de múltiplas oscilações detectadas.

A utilização do método implica na aplicação de um pré-processamento que elimine fortes distúrbios e tendências não estacionárias (passo 1). No artigo não foi explicitado

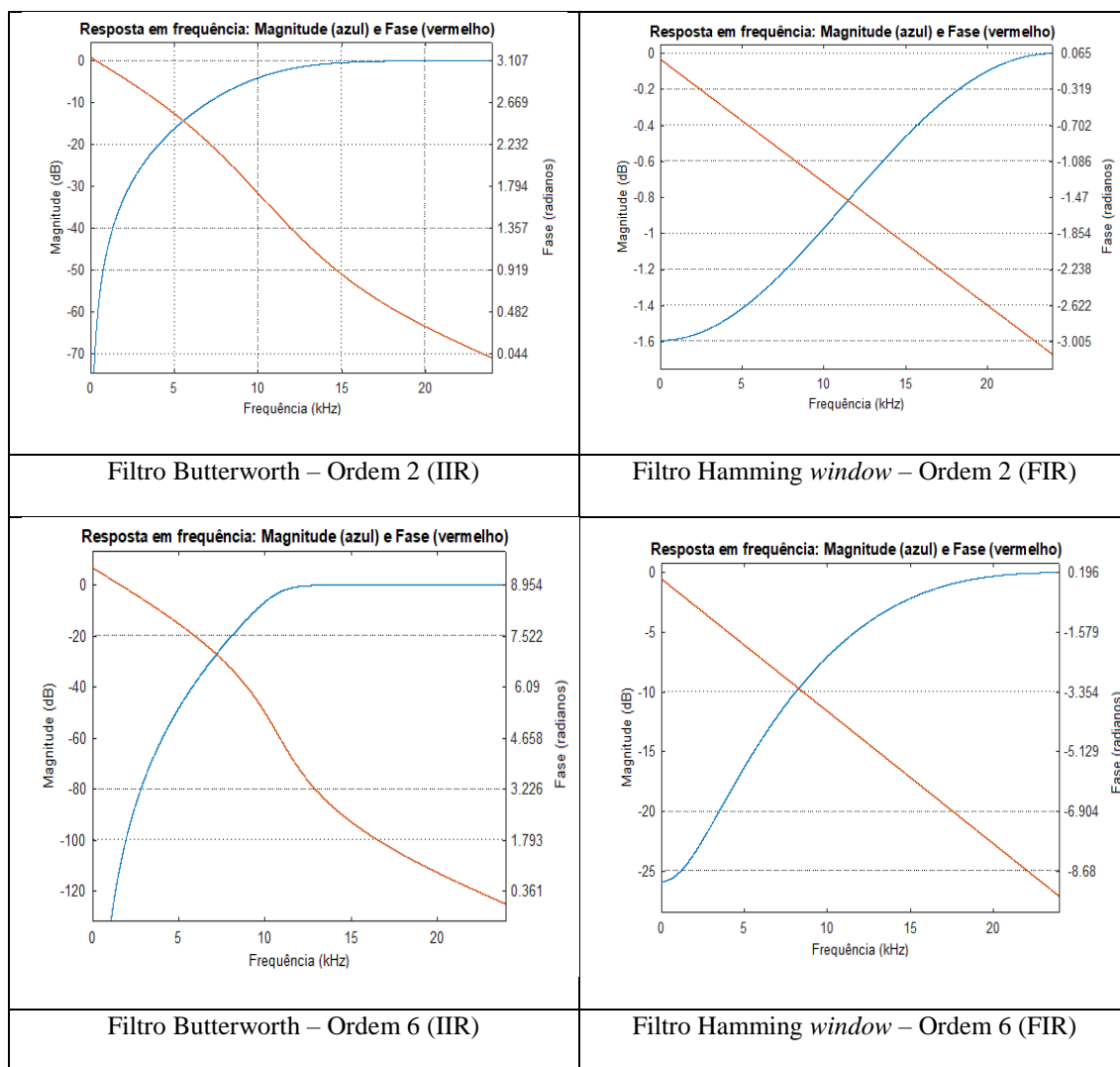
qual metodologia foi usada para esta etapa do algoritmo, por esse motivo será feita uma discussão sobre filtros.

Os filtros digitais podem ser classificados em duas categorias de acordo com a implementação: filtros com resposta ao impulso de duração infinita (*Infinite Impulse Response* - IIR) que são implementados de forma recursiva ou filtros com resposta ao impulso de duração finita (*Finite Impulse Response* - FIR) que são implementados através da convolução do sinal de entrada com a resposta ao impulso do filtro (SMITH, 1999). Não existe uma resposta única e definitiva para qual o filtro deve-se usar para cada caso.

Prosseguindo com a comparação entre os filtros FIR e IIR, cada um possui vantagens e desvantagens em relação ao outro. Sabe-se que os filtros FIR são estáveis, enquanto os filtros IIR tem fase não linear devido a adição de polos no sistema, o que causa distorções na fase, além disso, os filtros FIR em geral possuem uma implementação mais simples do que os filtros IIR. Contudo, se a aplicação não possuir restrições quanto a distorção de fase, os filtros IIR podem ser utilizados, tomadas as devidas precauções de verificação de estabilidade do filtro, que pode mudar de acordo com a seleção de parâmetros. A vantagem dos filtros IIR em relação aos filtros FIR é que em geral possuem uma ordem baixa e produzem bons resultados, reduzindo assim o esforço computacional se comparado aos filtros FIR que, para atingir um mesmo resultado, necessitam de filtros com ordem maior, aumentando o esforço computacional necessário.

A Figura 4 mostra um comparativo da resposta em frequência para dois exemplos de filtros (um filtro FIR e outro IIR). As curvas em azul representam o gráfico de amplitude da resposta em frequência, enquanto a curva vermelha representa o gráfico da fase. Conforme já foi abordado no parágrafo anterior, nota-se nos gráficos que a fase dos filtros IIR possui uma resposta não linear enquanto a fase dos filtros FIR possui uma resposta linear. Observa-se ainda que a não linearidade do gráfico de fase do filtro IIR se intensifica com o aumento da ordem, o que justifica a utilização de uma ordem menor para gráficos IIR, sempre que possível, além do esforço computacional reduzido.

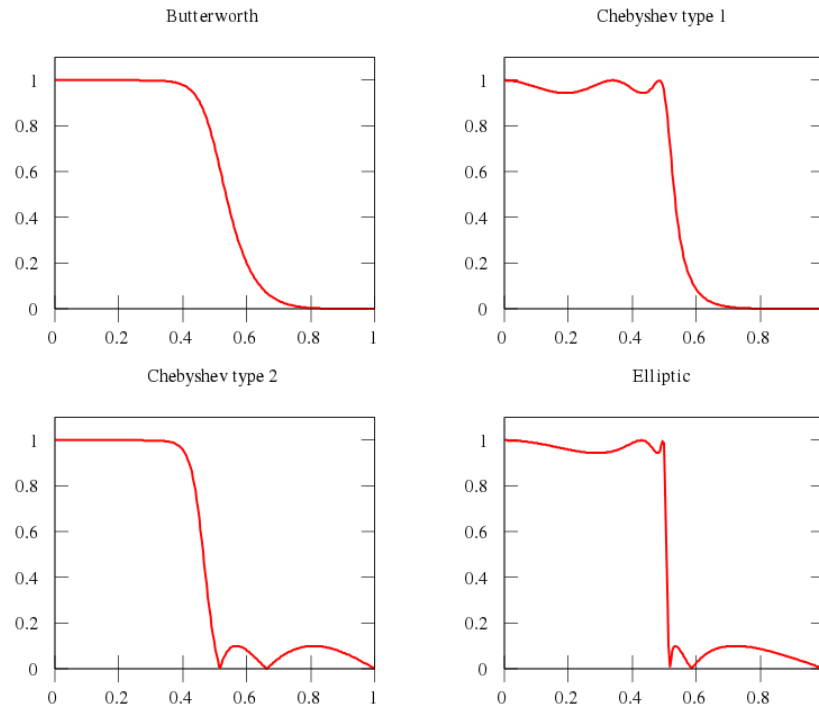
Figura 4 – Comparativo da resposta em frequência de filtros passa-altas IIR (esquerda) e FIR (direita) de ordens 2 e 6



Fonte: O próprio autor

Entre os filtros IIR podem ser citados os filtros de Butterworth, Chebyshev dos tipos I e II e o filtro elíptico. Uma comparação entre esses filtros pode ser feita analisando a resposta em frequência dos mesmos, Figura 5. Pode-se verificar que, dentre os filtros IIR o filtro de Butterworth possui vantagem por não possuir efeito *ripple* (distorção) nas bandas passante e rejeitada, porém possui o transitório mais lento quando comparado aos outros. Já os outros filtros citados possuem um transitório mais rápido com a desvantagem de possuírem *ripple* nas bandas. Portanto, conclui-se que, para uma aplicação onde as frequências que vão ser eliminadas do sinal estão distantes das frequências a serem mantidas, uma solução viável entre os filtros FIR é o Butterworth.

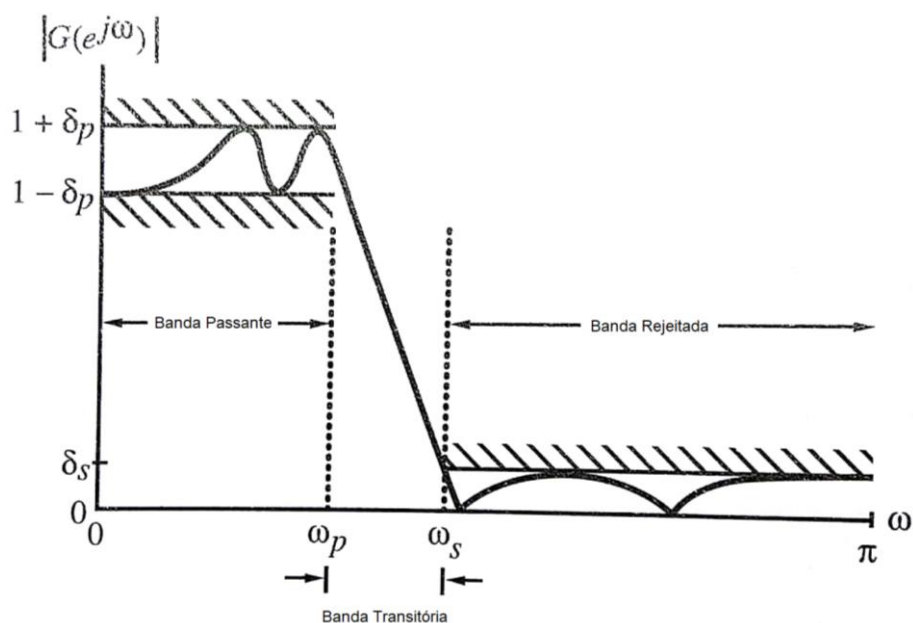
Figura 5 – Comparativo resposta em frequência de filtros IIR



Fonte: DAMATO, 2018

Diante das vantagens do filtro de Butterworth em relação aos outros filtros IIR, além de ter apenas 2 parâmetros para serem definidos (frequência de corte e ordem do filtro), o filtro de Butterworth foi a opção selecionada para este trabalho. A metodologia para obtenção dos parâmetros do filtro foi detalhada em Rabiner e Gold (1975). A ordem do filtro deve ser a menor possível que satisfaça às seguintes especificações de amplitude da resposta em frequência da função de transferência do filtro (Figura 6): frequência limite da banda passante (ω_p), frequência limite da banda rejeitada (ω_s), *ripple* da banda passante (R_p) e atenuação da banda rejeitada (R_s). As variáveis R_p e R_s são obtidas, respectivamente, a partir dos desvios δ_p e δ_s . A frequência de corte também é obtida através dos requisitos citados.

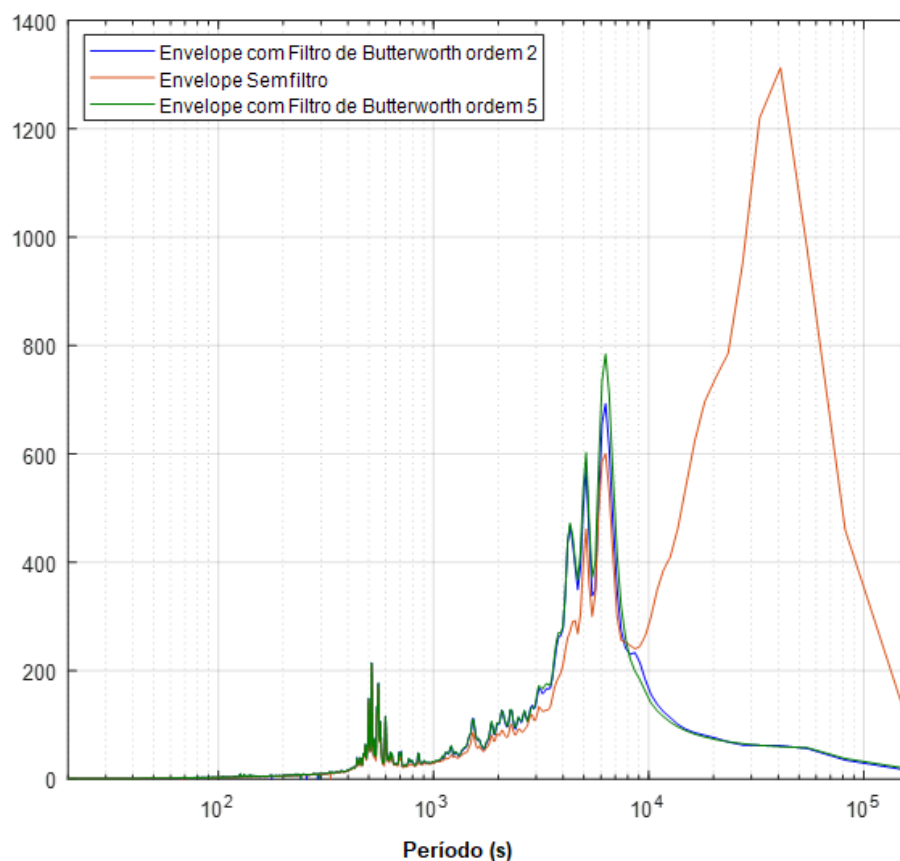
Figura 6 – Especificação de filtros do tipo FIR



Fonte: (MITRA, 2001)

Dessa forma, a ordem obtida para o filtro foi 2, que alcançou resultados desejados (Figura 7). O aumento da ordem apenas aumentou a complexidade e esforço computacional sem um ganho de desempenho que justificasse esse aumento. Além disso, conforme pode ser verificado na Figura 4, uma menor ordem apresenta uma menor distorção na fase devido a quantidade menor de polos adicionados.

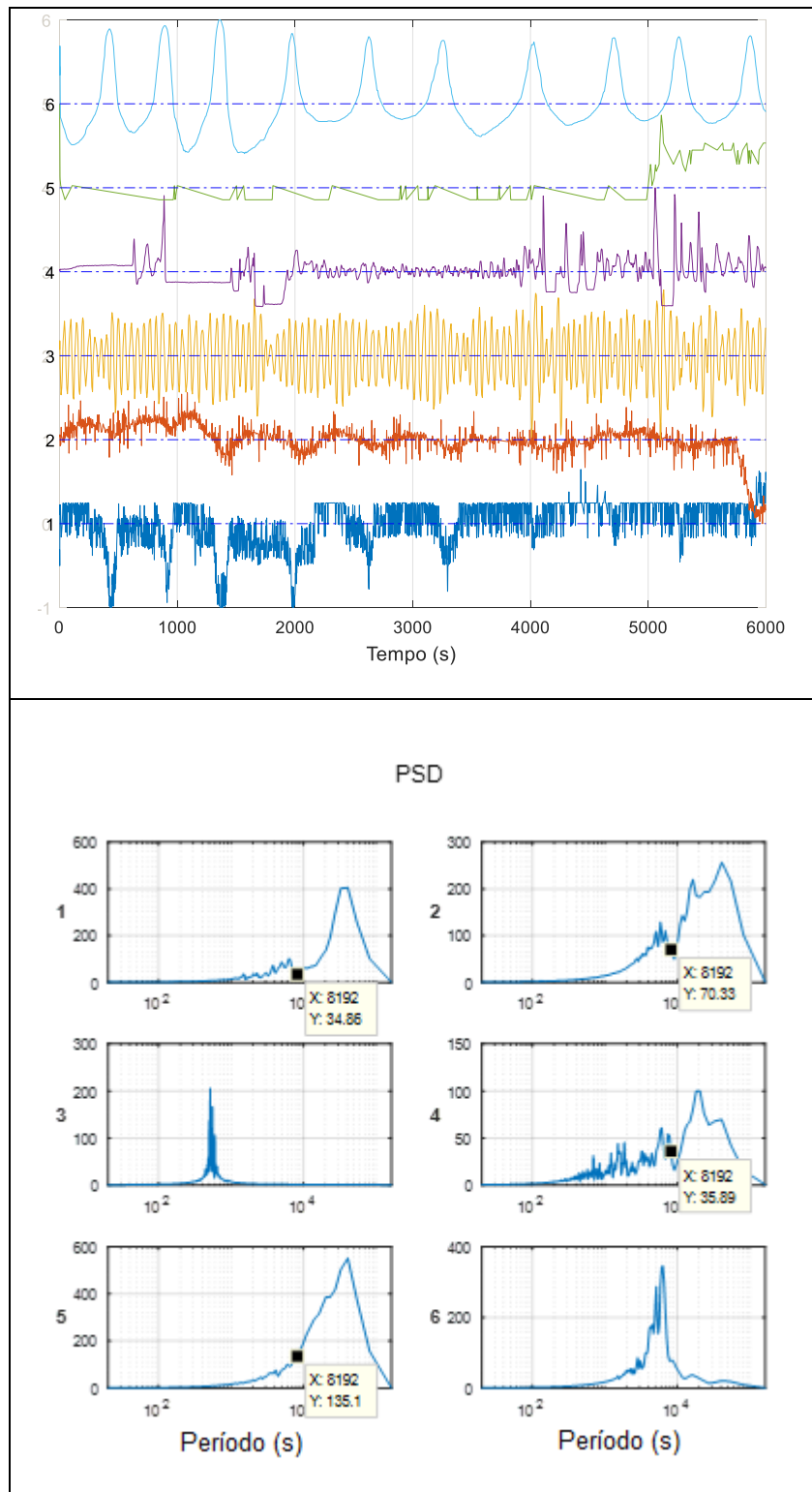
Figura 7 – Envelope espectral calculado sem filtragem e utilizando filtro de Butterworth de ordens 2 e 5



Fonte: O próprio autor

Em relação a frequência de corte, existem diferentes tipos de dados e possíveis aplicações para os algoritmos de detecção de oscilações. Diferentes tipos de processos apresentam diferentes características em relação as oscilações, variando desde oscilações de baixa a oscilações de alta frequência. No caso dos dados industriais utilizados no presente trabalho, foi identificado em conjunto com especialistas da indústria que não havia interesse em analisar baixas frequências. Baseando-se nos requisitos passados pelos especialistas da indústria e utilizando o procedimento de projeto do filtro descrito nesta seção, foi definida a frequência de corte igual a $1,25 \times 10^{-4}$ Hz (período de 8000s). Ao investigar o espectro de diversos processos de produção de óleo e gás notou-se que as componentes de frequência com período maior que 8000s estavam associados a distúrbios ou oscilações irregulares, que não contribuía para as análises (Figura 8). Fica claro na figura que o sinal 3, por exemplo, não possui distúrbios, mas oscilações bem definidas e regulares, o que pode ser notado tanto no sinal no domínio do tempo quanto em sua PSD, enquanto em outros sinais como o sinal 1, fica clara a presença de distúrbios.

Figura 8 – Sinais com distúrbios e suas respectivas PSD



Fonte: O próprio autor

3.1.2 Segmentação de dados

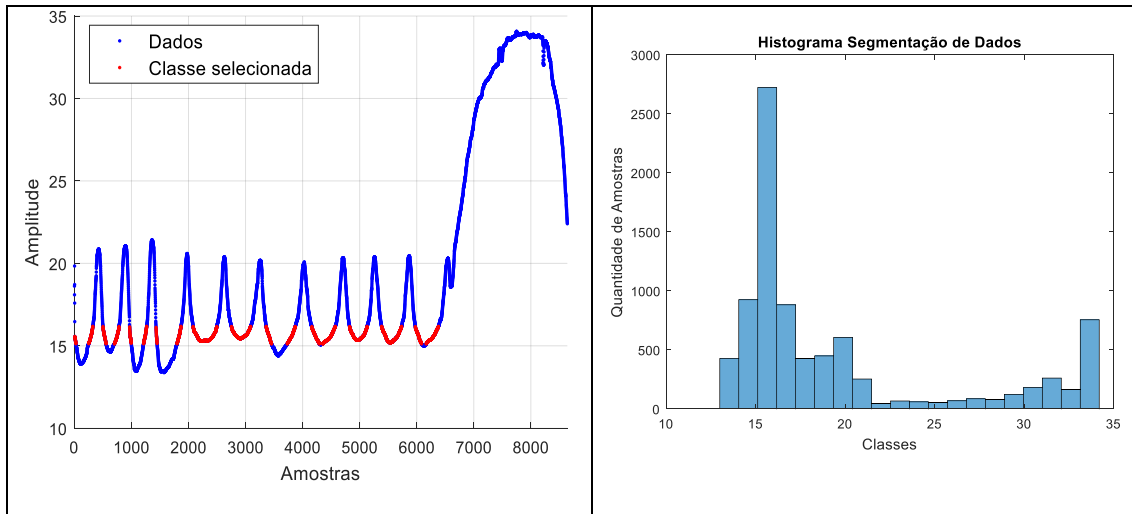
Uma outra abordagem que pode ser utilizada a fim de eliminar distúrbios dos sinais é a segmentação de dados. Essa metodologia é implementada através de um algoritmo com o objetivo de remover segmentos de dados que possuam distúrbios, utilizando propriedades estatísticas dos dados como a média e o desvio padrão (VIEIRA JUNIOR, 2016).

Seguem abaixo os passos da metodologia desenvolvida para eliminar os segmentos de dados com distúrbio e *outliers*.

1. Cálculo do histograma dos dados;
2. Seleção da classe que contém o maior número de amostras (conjunto de dados selecionados inicialmente);
3. Inclusão das classes na vizinhança no conjunto selecionado caso o número de ocorrências seja maior que um limiar baseado na quantidade total de dados;
4. Definir os limites mínimo e máximo das classes selecionadas e determinar o conjunto dentro dos limites (conjunto selecionado);
5. Calcular a média e desvio padrão do conjunto selecionado;
6. Recalcular os limites mínimo e máximo utilizando a média e desvio padrão calculados no passo 5;
7. Recalcular o conjunto selecionado dentro dos novos limites;
8. Encontrar descontinuidades no conjunto selecionado e manter apenas o maior segmento sem descontinuidades.

O histograma é uma ferramenta gráfica utilizada normalmente em aplicações estatísticas com o objetivo de representar a distribuição de frequências de um conjunto de dados. Nesse caso foi utilizado nos passos 1 e 2 da metodologia proposta para separar os dados em uma quantidade de *bins* ou classes de acordo com o número de ocorrências. No passo 2, a classe com maior frequência é definida como conjunto de dados inicialmente selecionado como ponto de partida. A Figura 9 mostra um exemplo de um sinal que foi dividido em 20 *bins* e o respectivo histograma.

Figura 9 – Primeira etapa da segmentação de dados: classe inicialmente selecionada (em vermelho) e histograma (direita)



Fonte: O próprio autor.

A Figura 9 mostra um exemplo de um conjunto de dados contendo oscilações e um distúrbio. O histograma com 20 *bins* mostra que a terceira barra do histograma representa a classe inicialmente selecionada (valores de amplitude próximos de 15), com maior número de elementos. Na cor vermelha, pode-se visualizar os elementos da classe inicialmente selecionada.

A partir da classe selecionada, é feita a verificação das classes próximas (ordem decrescente de quantidade de elementos) para definir o conjunto selecionado. O critério de seleção das classes próximas é o limiar definido pelo usuário (*default* = 2,5% da quantidade total de dados). No limiar definido como exemplo a classe próxima é incluída no conjunto selecionado se tiver pelo menos 2,5% da quantidade total de dados.

Após a seleção inicial dos dados e toda vez que uma nova classe é incluída no conjunto selecionado de dados, devem ser calculados os limites (valores mínimo e máximo) da classe. Para isso é utilizada a expressão da Equação (13), onde μ_{xs} é a média do conjunto de dados selecionado, σ_{xs} é o desvio padrão do conjunto de dados selecionado e p o fator multiplicativo do desvio padrão que funciona de forma similar ao intervalo de confiança.

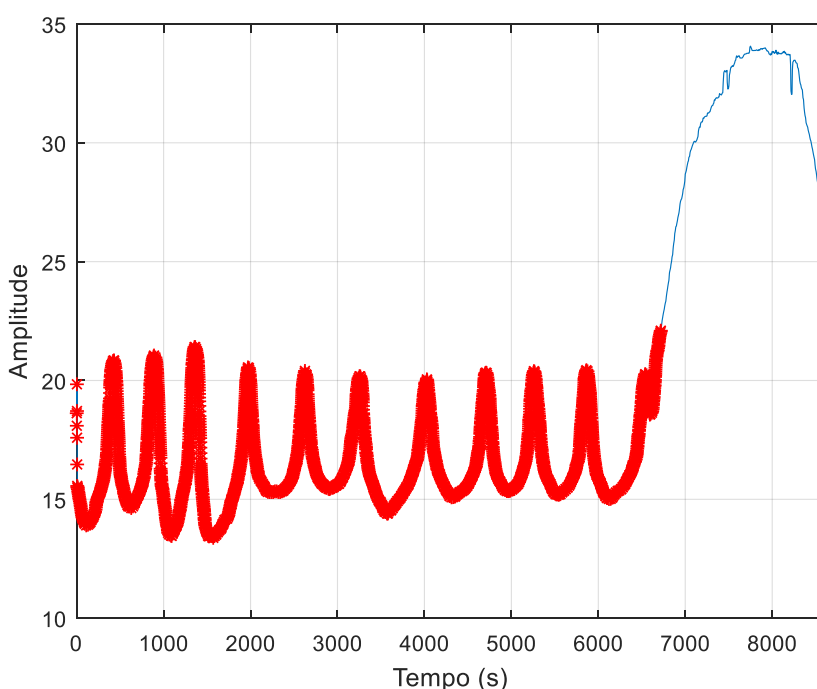
$$\begin{aligned} x_{min} &= \mu_{xs} - p\sigma_{xs} \\ x_{max} &= \mu_{xs} + p\sigma_{xs} \end{aligned} \quad (13)$$

Dessa forma o conjunto de dados selecionado pode ser definido verificando se os valores dos dados estão no intervalo $[x_{min}, x_{max}]$. O procedimento de cálculo da média

e de desvio, bem como do intervalo do conjunto de dados selecionados é repetido à medida que se verifica a inclusão de novas classes ao conjunto de dados.

A última etapa do algoritmo é a verificação de descontinuidades no conjunto de dados selecionado. Um parâmetro definido pelo usuário que representa o número de pontos entre dois segmentos sucessivos é utilizado para verificar descontinuidades. Caso o número de pontos entre segmentos supere o limiar estabelecido pelo parâmetro a descontinuidade é identificada e os segmentos são unidos. Após o procedimento completo executado pelo algoritmo o resultado final pode ser verificado na Figura 10.

Figura 10 – Resultado da segmentação de dados (dados selecionados em vermelho)



Fonte: O próprio autor.

O algoritmo de segmentação de dados implementado seguindo a metodologia acima elimina a necessidade de projeto de filtros e definição de parâmetros, que podem dificultar o tratamento de dados. O algoritmo de segmentação de dados reduz significativamente a dificuldade em definir parâmetros, porém não elimina completamente esta necessidade. Existem alguns parâmetros do algoritmo que podem ser definidos de acordo com o objetivo que se deseja alcançar e características dos dados sendo analisados. Por exemplo, alterando-se o desvio padrão, um conjunto menor de dados tende a ser selecionado, ou seja, os resultados passam a ser mais restritos. Os parâmetros do algoritmo são: número de *bins* ou classes do histograma (*nb*), número de pontos sucessivos entre dois

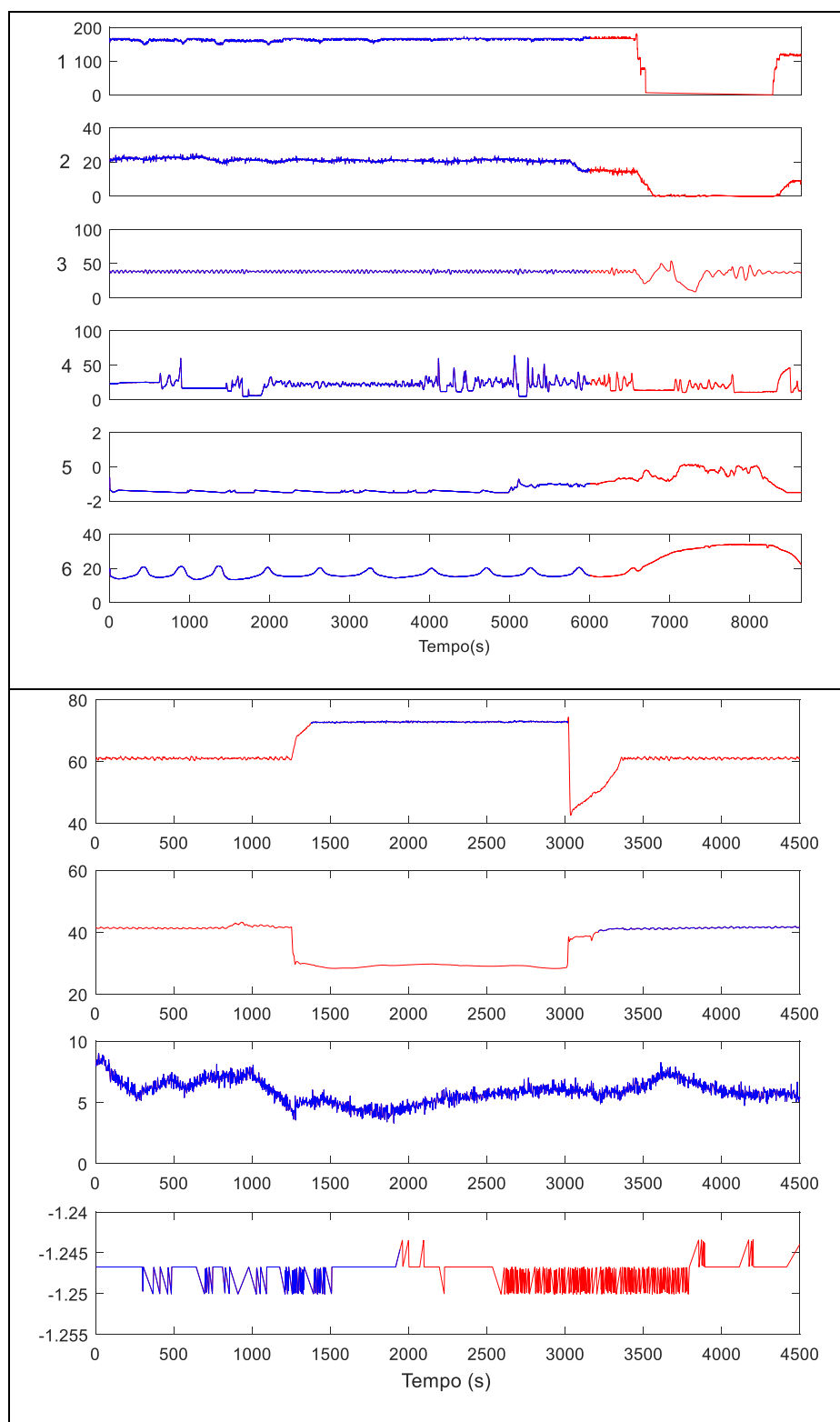
seguimentos (nd), multiplicador do desvio padrão do intervalo de confiança (ndp) e o limiar percentual dos dados a serem incluídos no conjunto inicialmente selecionado (limiar).

A grande vantagem de se utilizar o algoritmo de segmentação de dados é a possibilidade de eliminar diferentes tipos de distúrbios com uma mesma ferramenta, sem a necessidade de utilização de filtros. Ao se utilizar filtros surge a necessidade de escolher qual o mais adequado para a aplicação e também definir os parâmetros e outros procedimentos relacionados ao projeto de filtros.

A desvantagem de utilizar a segmentação de dados surge de um dos requisitos do envelope espectral. O cálculo do envelope espectral apenas pode ser feito se os diferentes sinais estão amostrados simultaneamente. O algoritmo de segmentação é executado de forma individual em cada sinal da matriz de dados e diferentes segmentos são retirados dos sinais, deixando os sinais com diferentes tamanhos, dessa forma, a segmentação de dados resolve muitos dos problemas relacionados a eliminação de distúrbios, porém, sua utilização só pode ser feita quando um mesmo segmento dos dados de todos sinais de interesse estiver disponível.

A Figura 11 mostra as duas situações relatadas no parágrafo anterior onde o conjunto de dados superior mostra em azul o segmento selecionado pelo algoritmo de segmentação de dados e em vermelho o segmento descartado. Nota-se que os distúrbios ocorrem em todos os sinais aproximadamente no mesmo instante, dessa forma a utilização do algoritmo é viável (para aplicar ao método do envelope espectral). No conjunto de dados da parte inferior da Figura 11, diferentes segmentos foram retirados dos dados, resultando em sinais com diferentes tamanhos e que não foram amostrados simultaneamente, impossibilitando a utilização do envelope espectral.

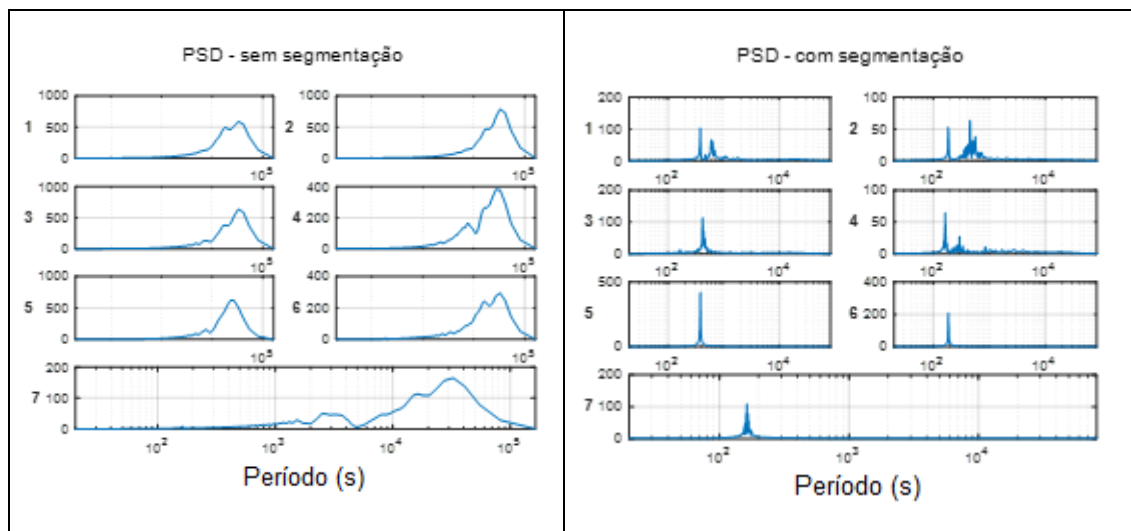
Figura 11 – Segmentação de dados aplicada a diferentes conjuntos de dados



Fonte: O próprio autor.

A Figura 12 mostra um comparativo da PSD de cada sinal antes e após a segmentação de dados, evidenciando a efetividade do pré-tratamento dos dados para eliminação de distúrbios.

Figura 12 – Comparativo entre a PSD calculada sem segmentação de dados e com a segmentação de dados



Fonte: O próprio autor.

3.1.3 Remoção de tendências lineares

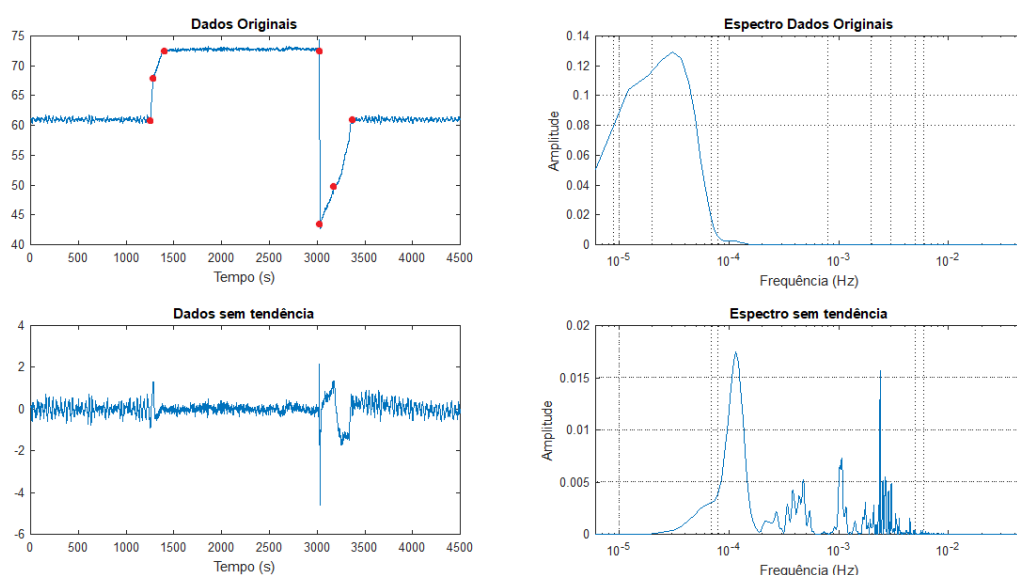
Como foi visto na seção anterior, existem situações em que é possível utilizar a segmentação de dados para remover tendências e outros distúrbios, e casos em que a remoção de distúrbios resulta em uma condição que impossibilita a utilização do envelope espectral. Para as situações em que não é possível utilizar a segmentação de dados, outros tratamentos devem ser empregados, entre eles a remoção de tendências lineares se apresenta com uma alternativa.

Tendências lineares são alterações comumente encontradas em séries temporais oriundas de diversas áreas, desde mercado financeiro até dados da indústria. Alguns dos sinais industriais utilizados nesse trabalho (Figura 3) apresentam esse tipo de comportamento, que tem uma influência negativa no espectro (distúrbio) e faz com que o sinal não possua mais característica estacionária.

Uma solução para o problema é a remoção de tendências lineares dos sinais. Esse procedimento consiste na seleção dos *breakpoints*, que são os pontos onde ocorre a mudança de comportamento dos dados, chamadas de tendências lineares (crescente ou

decrecente), por exemplo, uma rampa (Figura 13). Em seguida deve ser feita a regressão linear para estimar uma reta adequada à tendência que os dados apresentam. Como as tendências lineares podem ser aproximadas por equações de primeira ordem a remoção de tendências torna-se uma tarefa simples. A maior dificuldade está na seleção dos *breakpoints*. Algumas vezes esse procedimento de seleção pode ser subjetivo e pode ser realizado manualmente ou através de algoritmos utilizando rotinas de programação baseados em soma cumulativa (BROWN, R. L.; DURBIN, J; EVANS, 1975), por exemplo. Alguns trabalhos (DAVIS; LEE; RODRIGUEZ-YAM, 2006; DURBIN, 1969; KILLICK; FEARNHEAD; ECKLEY, 2012) apresentaram outras propostas para identificar pontos em que as características estatísticas dos dados mudam. De posse do modelo da tendência, é feita a subtração do sinal original pelo modelo (obtido através da regressão), restando assim apenas o sinal sem tendências (SHUMWAY; STOFFER, 2006, 2011).

Figura 13 – Resultado da aplicação de remoção de tendências lineares (PSD à direita)



Fonte: O próprio autor.

Na Figura 13 nota-se que a remoção de tendências lineares utilizando 7 *breakpoints* eliminou os maiores distúrbios no espectro, porém nota-se que variações abruptas verticais do sinal não foram completamente removidas, necessitando descartar o segmento em que esse tipo de alteração ocorre no sinal para eliminá-las completamente. Ainda assim, verificou-se uma melhora significativa no espectro do sinal.

Existem tendências com comportamento estocástico e também tendências determinísticas com comportamento não linear, entretanto essas tendências não estão no

escopo deste trabalho pois são casos mais específicos e que requerem um tratamento especial para sua remoção dos sinais. Uma alternativa para remover tendências tanto lineares quanto não lineares é a utilização de *splines*, que são curvas definidas por pontos (DE BOOR, 2001). As *splines* podem ser utilizadas de forma semelhante a regressão linear para obter uma estimação da curva original a partir de pontos amostrados. Dessa forma, o distúrbio pode ser removido através da diferença entre a curva original e a curva estimada (KUCHARCZYK; WYŁOMAŃSKA; ZIMROZ, 2017).

Uma observação que deve ser feita é que a utilização de remoção de tendências lineares não elimina a necessidade de utilizar um filtro. Um mesmo sinal pode conter tendências lineares e distúrbios oscilatórios de baixa frequência que serão mantidos mesmo após a remoção dessas tendências. Cada tratamento possui o objetivo de eliminar um tipo específico de distúrbio. Comparando a remoção de tendências lineares com a segmentação de dados a vantagem clara já foi citada no início desta seção: a remoção de tendências lineares não descarta amostras dos dados, possibilitando a utilização do envelope espectral em quaisquer dados que foram tratados por essa metodologia. A desvantagem é a necessidade, em algumas situações, de utilização de filtro em conjunto com a remoção de tendências lineares, o que pode também acontecer na segmentação de dados já que, dependendo da amplitude do distúrbio de baixa frequência, este não será descartado pela segmentação de dados, tornando necessária também a utilização de filtro em conjunto com esta metodologia.

3.2 Impacto das oscilações nos sinais analisados

Muitas das análises realizadas nesse trabalho tratam ao mesmo tempo diversos sinais de um mesmo processo. Um exemplo são os dados de sensores de um processo de compressão de gás, um dos muitos processos de uma unidade de produção de óleo e gás. Esse processo de compressão de gás possui 21 sinais coletados de sensores variados como, por exemplo, temperatura, pressão e nível. Não há interesse em detectar oscilações que tenham pequena amplitude, pois causam pouco impacto nas malhas de controle, portanto, sinais com estas características devem ser eliminados das análises. Além disso, o custo dessas variações no sinal é proporcional a variabilidade do sinal (SHINSKEY, 1990), o que reforça a importância de se analisar essa característica mantendo ou descartando sinais antes mesmo de se calcular o espectro e outras análises subsequentes, poupando assim esforço computacional e tempo de análise.

Uma maneira de estimar o impacto dos sinais é através de medidas de dispersão dos dados. Entre as medidas de dispersão podem ser citadas a variância, o desvio padrão e o coeficiente de variação. O desvio padrão amostral s e a variância amostral s^2 podem ser obtidos através da Equação (14), onde x_i é a i -ésima amostra dos dados, μ_x a média e n o número de amostras do vetor de dados sendo analisado (KOTTEGODA; ROSSO, 2008).

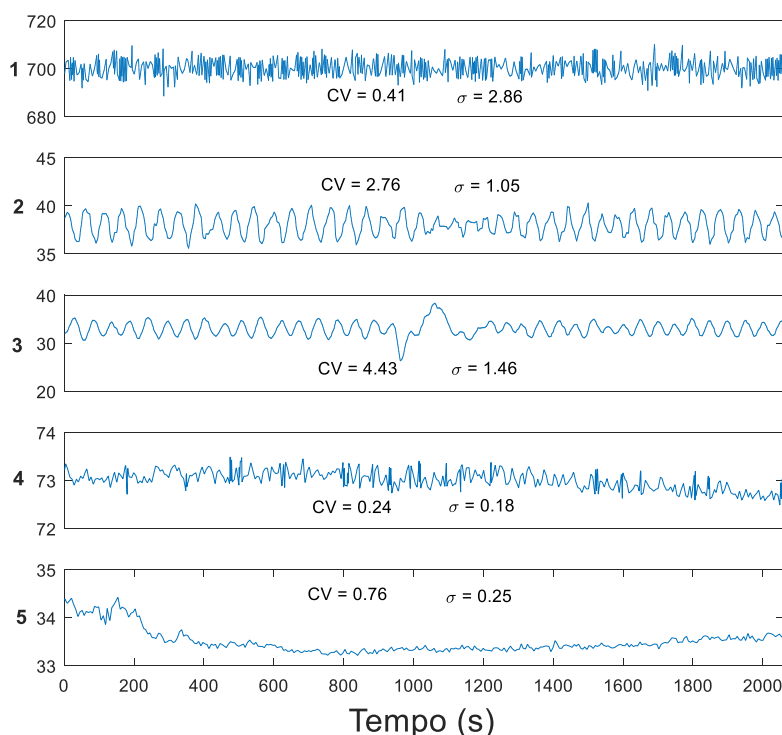
$$s = \sqrt{s^2} = \sqrt{\sum_{i=1}^n \frac{(x_i - \mu_x)^2}{n - 1}} \quad (14)$$

A variância é uma medida usual de dispersão da variável analisada, contudo existem outras medidas de dispersão como o coeficiente de variação (CV) dado por

$$CV = \frac{s}{\mu_x} \quad (15)$$

O coeficiente de variação foi a medida de dispersão escolhida para analisar o impacto das oscilações por uma escolha dos usuários do algoritmo (pesquisadores e equipe de operação), pois apresenta algumas vantagens em relação a variância ou desvio padrão. A primeira vantagem em se utilizar o coeficiente de variação está na possibilidade de se comparar sinais com diferentes unidades de medida (pressão em Pascal ou bar) e diferentes naturezas (variáveis de pressão, vazão, nível, temperatura), mudanças em unidades de medida não influenciam a análise. Além disso, uma outra vantagem que pode ser citada é o fato de ser uma medida adimensional que pode ser expressa em porcentagem, essa vantagem pode ser vista no conjunto de dados da Figura 14 em que são mostrados dados não normalizados e seus respectivos CV e desvio padrão. Nota-se que nos sinais 4 e 5 não há grande vantagem em se utilizar o CV em relação ao desvio padrão, porém nos outros sinais, principalmente no sinal 1 nota-se uma grande diferença entre as medidas. O sinal 1 possui um valor médio alto (em torno de 700), por esse motivo, analisar apenas o desvio padrão para descartar sinais de diferentes dimensões pode levar a um resultado diferente. Nessas situações, a melhor opção é utilizar o CV como medida de impacto. Porém, se os dados estiverem normalizados, o valor da média vai ser muito baixo (próximo de zero), fazendo com que o valor do CV segundo a Equação (15) se aproxime de um valor muito alto (tende a infinito). Nessas situações o melhor é utilizar apenas o desvio padrão como medida de impacto.

Figura 14 – Sinais no domínio do tempo com respectivos valores de desvio padrão e CV

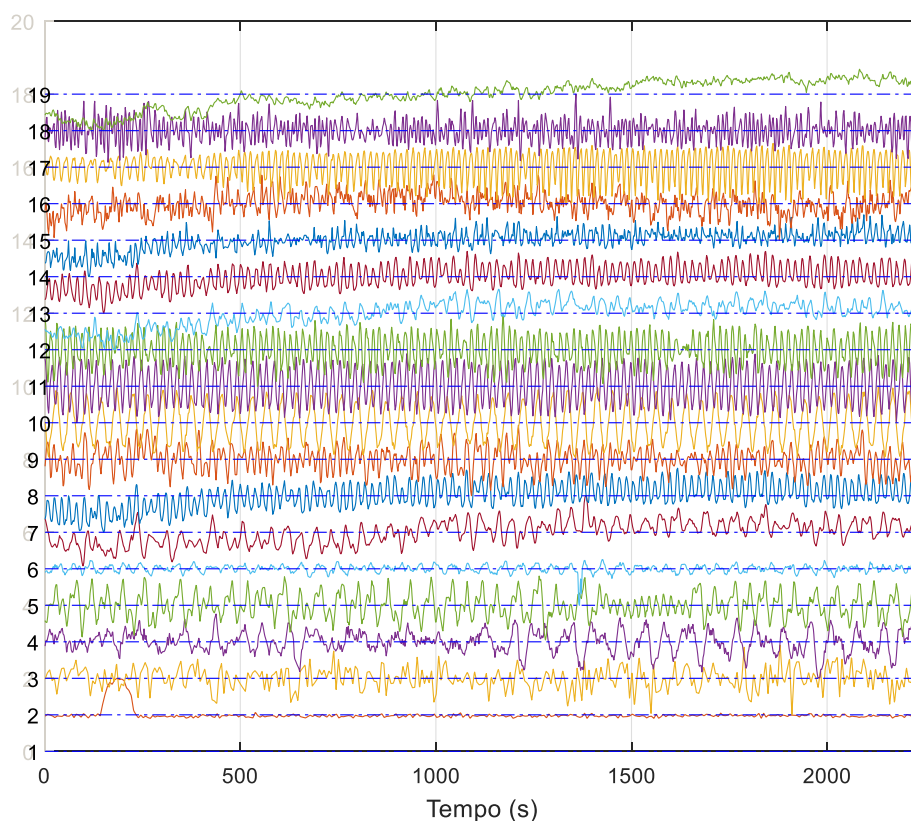


Fonte: O próprio autor.

Uma outra observação que pode ser feita em relação a seleção de sinais utilizando o CV (Figura 14) é que os sinais que apresentam menor CV (1, 4 e 5) são constituídos majoritariamente por ruídos, nem sequer apresentam oscilações regulares. Dessa forma, utilizar um limiar de 0,8%, por exemplo, descartaria esses sinais da análise, pois não há necessidade de investigar esses sinais que não possuem oscilações.

A Figura 15 mostra um conjunto de dados com 19 sinais obtidos a partir de um processo de compressão de gás. O conjunto de dados já passou por um tratamento de segmentação de dados.

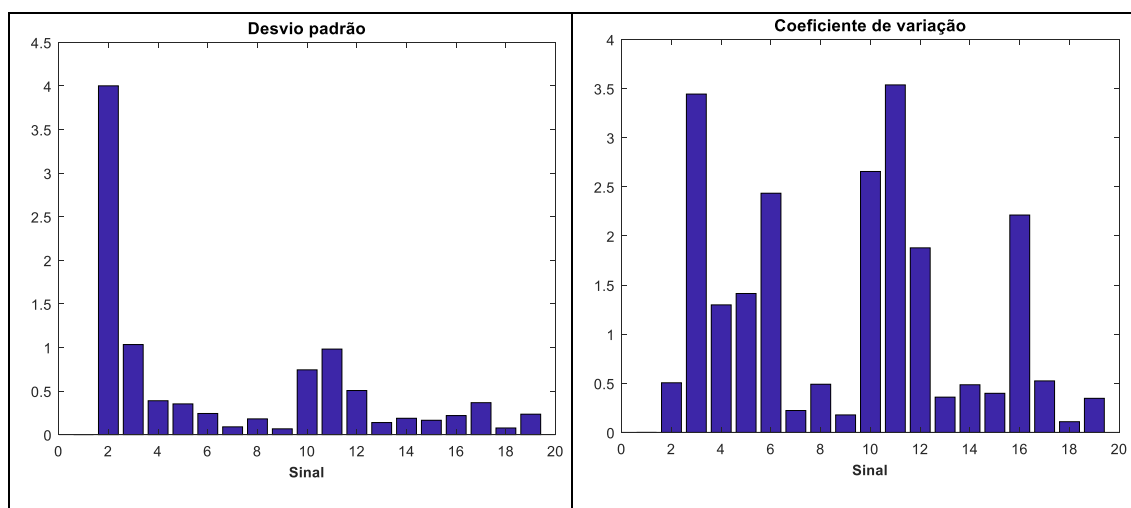
Figura 15 – Dados que passaram por tratamento (segmentação de dados)



Fonte: O próprio autor.

Após calcular o desvio padrão e o CV de cada sinal mostrado na Figura 15, o resultado é apresentado nos gráficos de barra da Figura 16. Nota-se no gráfico da direita que, ao se utilizar um limiar de 0,8%, por exemplo, seriam descartados muitos sinais das análises subsequentes (1, 2, 7, 8, 9, 13, 14, 15, 17, 18 e 19). Novamente, muitos deles representam sinais que nem sequer possuem oscilações regulares, como os sinais 1, 2 e 3. Comparando essa análise com o gráfico da esquerda (desvio padrão), pode-se verificar que um resultado muito diferente seria obtido ao aplicar limiares. Os sinais com maior desvio padrão são os sinais 1 e 2, exemplos de sinais que não possuem oscilações, mas possuem ruído.

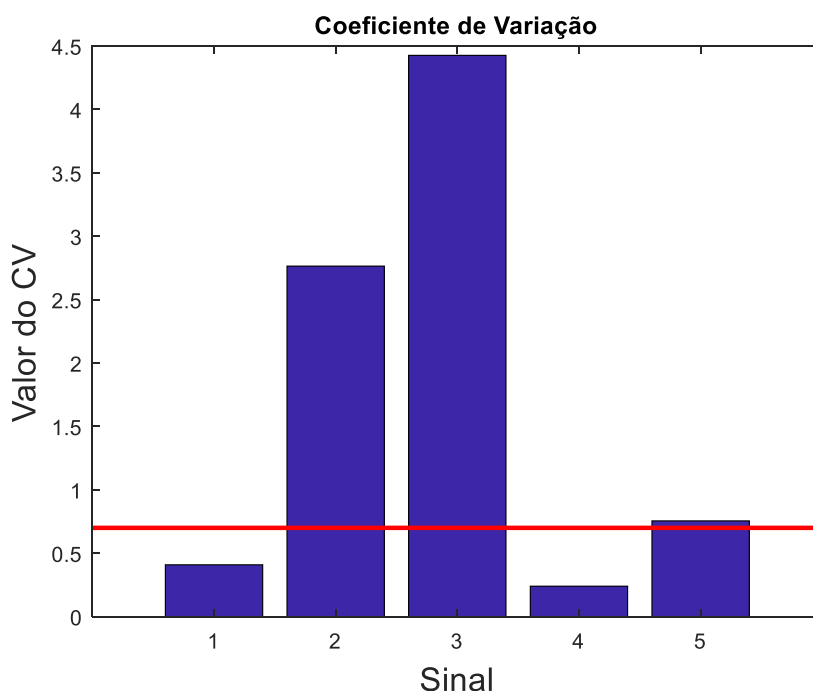
Figura 16 – Desvio padrão e coeficiente de variação de sinais analisados



Fonte: O próprio autor.

O limiar do CV deve ser discutido com a equipe de operação, que tem um melhor conhecimento do processo e do que se considera uma oscilação relevante, dessa forma o limiar pode ser definido. Esse valor limite pode variar de acordo com o tipo de malha de controle que está sendo analisada. Em malhas de nível tipo pulmão, por exemplo, um alto valor de variabilidade é aceito. Mesmo dentro de um mesmo processo, diferentes malhas admitem diferentes valores de CV. Por esse motivo, o ideal é que o limiar do CV seja analisado, discutido e definido especificamente para cada malha. A Figura 17 mostra o mesmo conjunto de dados da Figura 14, situação em que foi definido, em conjunto com a equipe de operação, que os sinais com CV inferior a 0,8% deveriam ser descartados das análises. Apesar desse limiar ter sido definido em conjunto com a equipe de operação, diferentes limiares podem ser definidos de acordo com o processo que está sendo analisado.

Figura 17 – Coeficiente de variação e o limiar de 0,7% aplicado aos sinais analisados



Fonte: O próprio autor.

Uma importante observação que deve ser feita em relação à utilização do CV para descartar sinais com baixo impacto das análises é que seu cálculo deve ser feito após o tratamento dos dados e de preferência em um segmento estacionário dos dados. É fundamental que isso seja feito pois mudanças de ponto de operação, tendências e fortes distúrbios alteram o valor do CV de um sinal. Como deseja-se analisar o impacto das oscilações e não apenas dos sinais, deve-se tomar este cuidado.

3.3 Seleção automática de intervalos de frequência

Em Jiang et al. (2007) e Wendt et al. (2000) os autores propõem a metodologia do envelope espectral e suas aplicações, e a utilização de recursos estatísticos como o teste de hipótese para a indicação de quais sinais apresentam oscilações. Em uma das aplicações apresentadas pelos autores, a metodologia proposta foi aplicada a dados industriais. Contudo, nas referências citadas, os autores utilizam a seleção manual dos picos de interesse para análise espectral. O procedimento manual de seleção de picos do envelope, referentes às oscilações, impõe uma grande limitação para a automatização da metodologia. Com o intuito de viabilizar a aplicação da metodologia de forma automática, esse trabalho propõe um método para realizar a seleção dos principais intervalos de frequência que serão utilizados para investigar a presença de oscilações.

Os passos do algoritmo de seleção automática de intervalos de frequência do envelope espectral são:

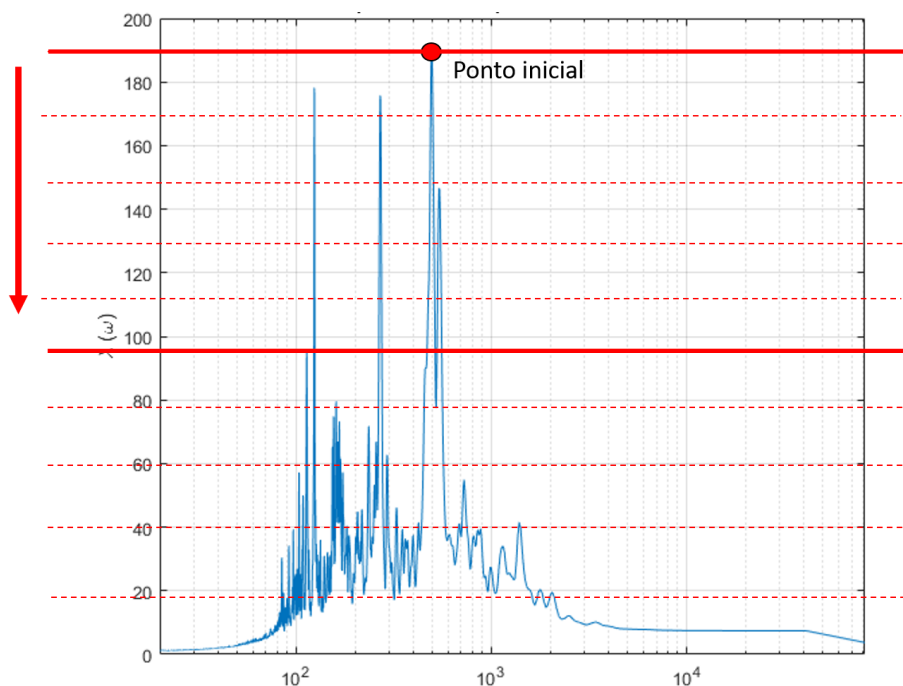
1. Selecionar os valores e a localização (índices) dos picos do envelope espectral;
2. Selecionar o maior pico e obter a máxima amplitude do envelope;
3. Fazer a varredura ou busca do limiar que corta o envelope e define os intervalos de frequência;
 - a. Iniciar varredura vertical (eixo das ordenadas) a partir do maior pico do envelope em direção a origem;
 - b. A cada iteração mover o limiar em uma unidade do divisor de busca e calcular o número de cruzamentos do envelope pelo novo limiar;
 - c. Caso o número de cruzamentos ultrapasse o dobro do número máximo de intervalos de frequência a busca chega ao fim e os cruzamentos são definidos como os intervalos de frequência;
4. Verificação de números ímpares de cruzamentos por zero;
5. Verificação da proximidade entre os intervalos, se forem próximos possivelmente representam uma mesma oscilação. Nesse caso, redefinir o intervalo como sendo as duas bordas externas dos dois intervalos.

Dois parâmetros devem ser definidos para a varredura executada no passo 3. O primeiro é o divisor de busca, por exemplo: 10% da máxima amplitude, que representa o quanto o limiar se desloca a cada iteração da varredura. O segundo parâmetro é o número máximo de intervalos de frequência que devem ser selecionados, por exemplo: 3 intervalos, referentes aos 3 maiores picos.

O algoritmo é iniciado com o procedimento de obtenção da amplitude dos picos do envelope espectral e seleção da maior amplitude associada a um pico. A maior amplitude é o ponto de partida da varredura que seleciona os intervalos de frequência. A Figura 18 ilustra o procedimento de varredura do passo 3 do algoritmo de seleção automática de intervalos de frequência, na qual um divisor é definido baseado na amplitude máxima e um limiar móvel que intercepta o envelope horizontalmente é movido a cada iteração. O quanto esse limiar se move a cada iteração é definido pelo divisor do passo 3-a. A busca é interrompida quando o limiar seleciona 3 intervalos de frequência ou quando se atinge um valor de amplitude mínimo igual a 10% da amplitude máxima. Por último, são feitas duas verificações: uma verificação necessária quando o vetor de cruzamentos possui quantidade ímpar, o que indica que um dos valores não pertence a um intervalo (par de

valores), valor que aparece nas extremidades devido a presença de distúrbios acentuados; A outra verificação é um procedimento de agrupamento necessário quando intervalos muito próximos são selecionados, intervalos que representam uma mesma oscilação.

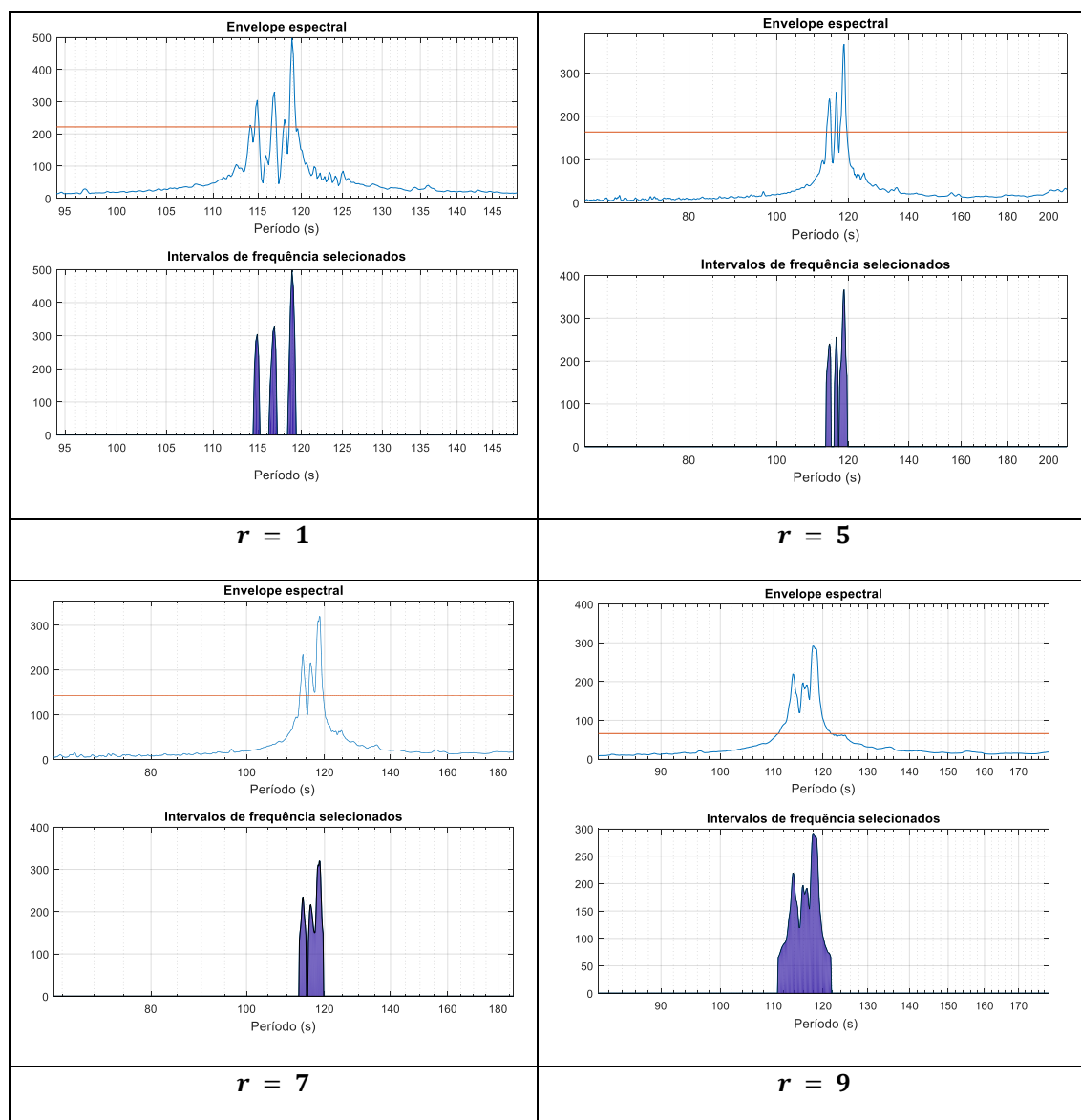
Figura 18 – Ilustração do algoritmo de seleção de intervalos



Fonte: O próprio autor.

Um aspecto que deve ser discutido em relação a utilização da seleção automática de intervalos de frequência no envelope espectral é o grau de suavização r do espectro (Equação (6)). Um grau de suavização baixo ($r = 1$) resulta em um espectro com maior quantidade de picos, sendo muitos deles picos estreitos e com baixa energia. Aumentando-se o grau de suavização ($r = 5$) nota-se que pequenos picos próximos aos picos principais passam a fazer parte dos picos principais e com isso uma menor quantidade de intervalos de frequência é selecionada para análise. Devido a este efeito, a suavização tem grande influência no resultado da seleção de intervalo de frequência, já que alguns picos podem deixar de existir ou passar a fazer parte de outros picos maiores, modificando assim o resultado da seleção. Esse efeito pode ser verificado na Figura 19 onde é possível visualizar o envelope espectral e os intervalos de frequência selecionados para diferentes valores de suavização.

Figura 19 – Resultado da seleção de intervalos de frequência para diferentes valores de suavização



Fonte: O próprio autor.

3.4 Proposta de utilização de um limiar de energia

Esta seção visa propor a utilização de um limiar de energia com o objetivo de descartar oscilações que estão relacionadas a sinais com muito ruído e/ou oscilações irregulares.

3.4.1 Cálculo da energia

Sabe-se pelo Teorema de Parseval (Equação (16)) que a energia total de um sinal no domínio do tempo, obtido através da soma quadrática do módulo do sinal é equivalente a soma da energia do sinal no domínio da frequência (KOOPMANS, 1995; STOICA, PETRE; MOSES, 2005), obtido através da área (integral) sob a curva de potência (PSD).

$$E_{TOTAL} = \sum_{t=-\infty}^{\infty} |y(t)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\omega) d\omega \quad (16)$$

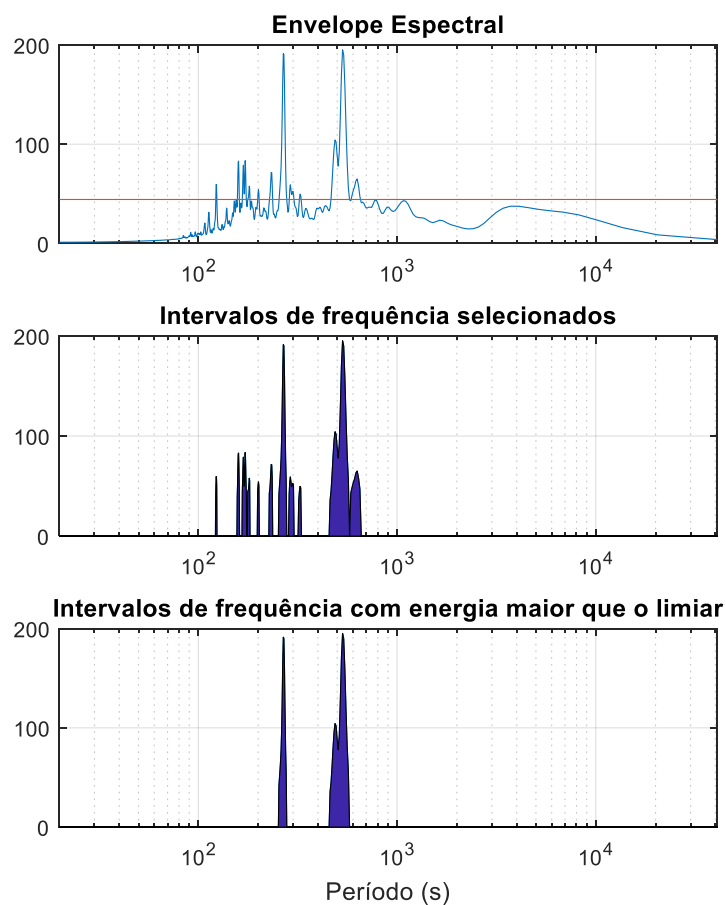
O envelope espectral é calculado utilizando uma ponderação $\beta(\omega)$ em cada frequência. Por esse motivo, a energia total do envelope espectral não pode ser calculada através da Equação (16). Contudo, a área sob a curva do envelope espectral fornece uma medida análoga à energia, que pode ser utilizada para descartar os intervalos de frequência com menor área que não são de interesse para as análises.

De posse dos intervalos de frequência obtidos na etapa anterior conforme foi descrito na seção 3.4 deste capítulo, a energia de uma oscilação pode ser obtida através da Equação (17). Novamente, as equações de energia definidas nas referências não são validas para o envelope espectral devido a ponderação aplicada nos cálculos, entretanto, de forma análoga, a área sob a curva dentro dos limites dos intervalos de frequência selecionados pode ser utilizada como uma medida para eliminar algumas oscilações.

$$E_{osci} = \sum_{\omega=\omega_1}^{\omega_2} \phi(\omega) \quad (17)$$

A área pode ser utilizada com uma medida percentual para cada oscilação, dividindo a área do intervalo de frequência pela área total sob a curva do envelope. Dessa forma, essa medida pode ser utilizada como um critério para descartar as oscilações irregulares ou que são compostas majoritariamente por ruídos. Um limiar percentual é utilizado para separar e eliminar intervalos de frequência de oscilações com baixa energia, ou seja, oscilações que não possuem grande impacto para a análise. A Figura 20 ilustra o procedimento completo desde a seleção de intervalos de frequência até o descarte de intervalos cuja medida da área está menor que o limiar percentual estabelecido. Ao final restaram intervalos de frequência de maior área correspondentes a oscilações regulares.

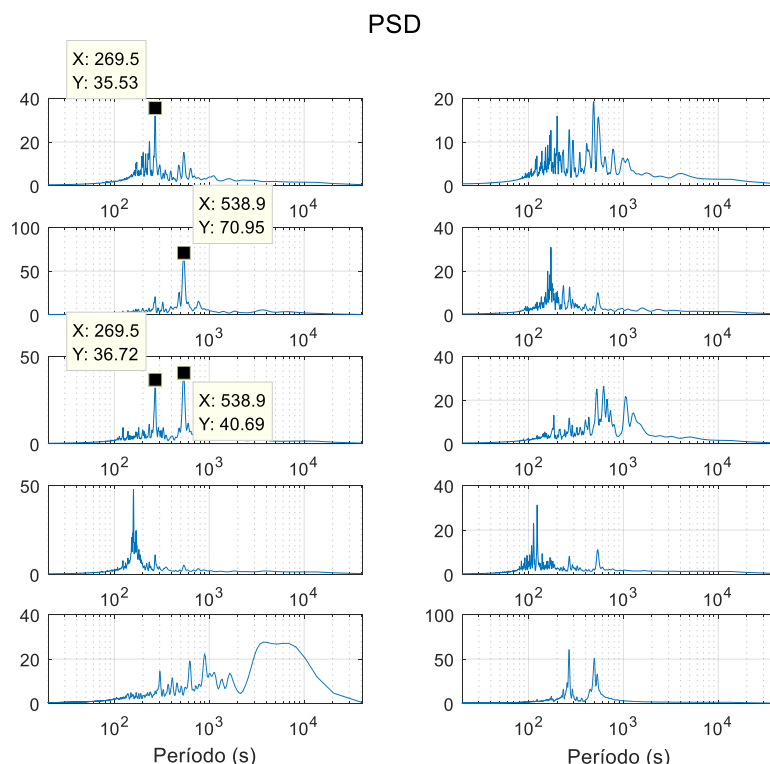
Figura 20 – Resultado da seleção automática de intervalos de frequência e da aplicação do limiar de energia para descartar intervalos



Fonte: O próprio autor.

Na Figura 21 é possível notar que as duas oscilações que permaneceram após o limiar ter sido aplicado são as que possuem maior amplitude e que se repetiram em mais de um sinal. As oscilações descartadas estão relacionadas a sinais com muito ruído e/ou oscilações irregulares.

Figura 21 – PSD dos sinais analisados após utilizar a seleção automática de intervalos de frequência e o descarte de intervalos de frequência



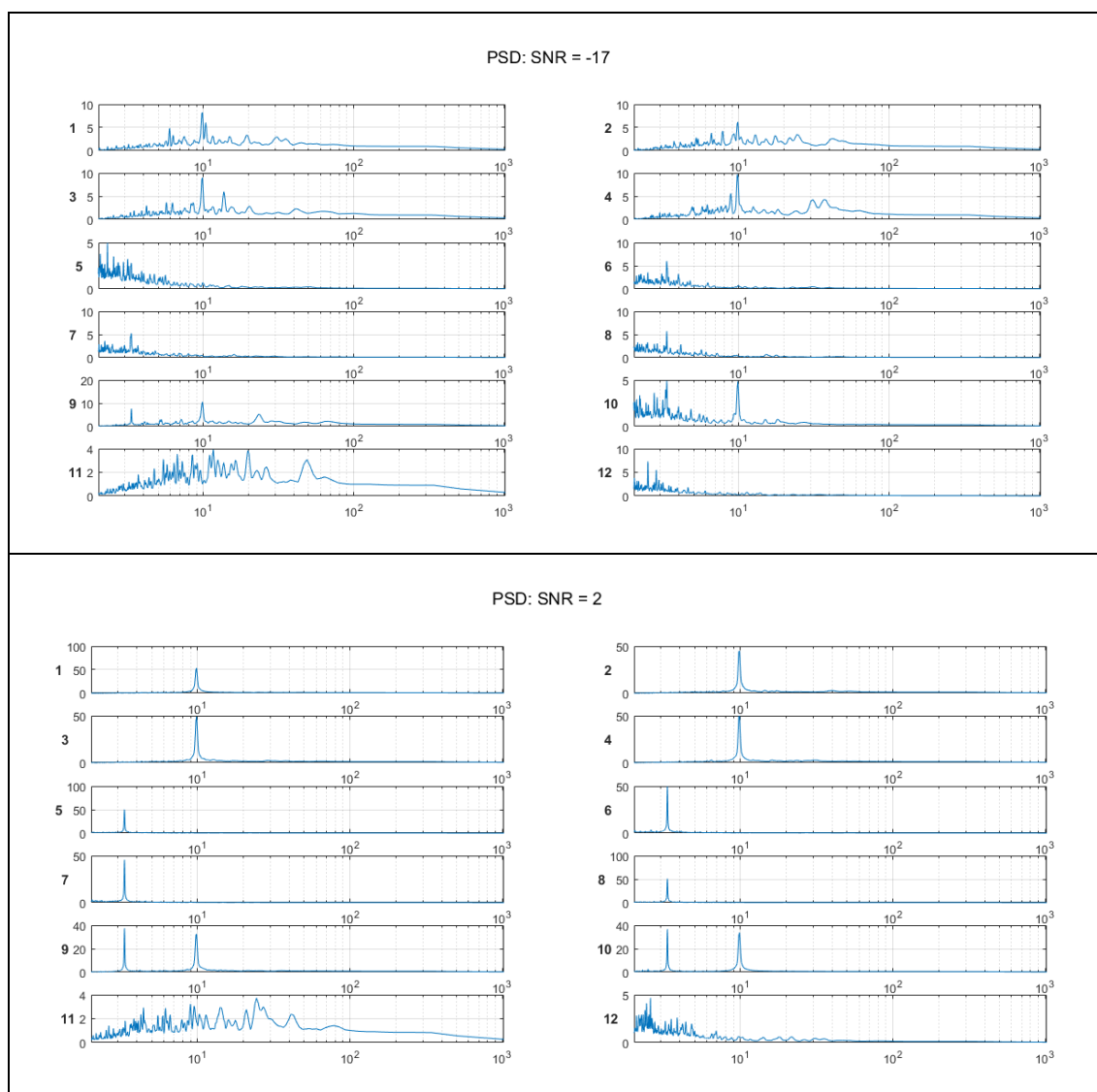
Fonte: O próprio autor.

3.4.2 Relação sinal/ruído e energia

Uma métrica comumente utilizada em diversas áreas da engenharia para avaliar a qualidade de sinais na presença de ruídos é a relação sinal-ruído, abreviada como S/N ou SNR. A relação sinal-ruído pode ser obtida através da divisão da potência do sinal sem qualquer ruído pela potência do sinal completo.

A presença de ruídos pode prejudicar a detecção de oscilações encobrindo oscilações presentes nos sinais. Um exemplo desta situação são os casos em que a oscilação tem amplitude próxima da amplitude dos ruídos, como pode ser visto na Figura 22, que mostra a PSD de 12 sinais simulados, os mesmos utilizados em Jiang et al. (2007), onde variou-se a amplitude do sinal mantendo-se fixa a amplitude do ruído. O primeiro gráfico mostra a PSD para um valor baixo de SNR enquanto o segundo gráfico mostra para um valor maior de SNR.

Figura 22 – PSD de sinais sintéticos para diferentes valores de SNR

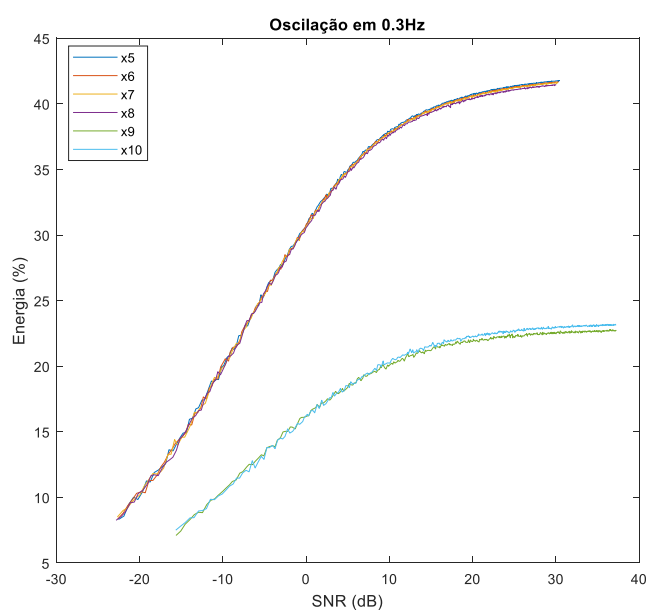


Fonte: O próprio autor.

A fim de investigar a influência da relação sinal-ruído na detecção de oscilações utilizando o envelope espectral foram utilizados dados sintéticos em que se pode variar os parâmetros da frequência das oscilações, amplitude do sinal e amplitude dos distúrbios, assim como os utilizados em Jiang et al. (2007). A seguinte simulação foi feita: a amplitude do ruído de todos os sinais foi fixada para todas as simulações realizadas, a amplitude dos sinais sem ruído foi variada, variando também a relação sinal-ruído do sinal completo. Para cada amplitude foram feitas 100 iterações do algoritmo e foi calculada a média das energias das oscilações nas 100 iterações realizadas, obtendo-se assim a média para cada valor de SNR.

O resultado das simulações pode ser observado no gráfico da Figura 23, os quais indicam a energia média das oscilações com relação ao SNR. Foram utilizados 10 sinais sintéticos (sinal oscilatório com ruídos, ambos simulados) com as seguintes características: os sinais 1 a 4 possuem uma oscilação de 0,1 Hz, os sinais 5 a 8 possuem uma oscilação de 0,3 Hz e os sinais 9 e 10 possuem as duas oscilações. Na Figura 23 apenas as curvas dos sinais contendo a oscilação em 0,3 Hz (sinais 5 a 10) são apresentadas. Algumas conclusões podem ser obtidas a partir das curvas mostradas. A primeira é que as curvas dos sinais 9 e 10, por possuírem 2 oscilações, possuem aproximadamente metade da energia das oscilações dos outros sinais. A segunda observação mostra que para sinais muito contaminados por ruídos (baixo SNR) a energia percentual é baixa também, e, de forma análoga à medida que se aumenta o SNR aumenta a energia percentual da oscilação. Nota-se também que a curva tem um comportamento aproximadamente linear até um SNR igual a 10 e para altos valores de SNR a energia começa a estabilizar em um valor próximo de 40%. A curva em questão pode ser utilizada como um guia para definir o limiar de energia. Um baixo SNR indica sinais com muito ruído, que não são de interesse para as análises. Segundo a curva, utilizar um limiar de 20%, por exemplo, descartaria sinais com SNR menor que -10, para os sinais em que se tem apenas 1 oscilação por sinal e descartaria a maioria dos sinais para os casos em que se tem 2 oscilações em um mesmo sinal, mantendo apenas poucas oscilações com SNR alto (acima de 20).

Figura 23 – Análise de energia e SNR



Fonte: O próprio autor.

4 Estudos de caso

O presente capítulo tem o objetivo de mostrar os resultados da aplicação da metodologia apresentada no capítulo 2 e das contribuições propostas no capítulo 3 em dados industriais, obtidos a partir de leituras de sensores de diferentes processos de uma unidade de produção de óleo e gás da Petrobrás.

Os algoritmos utilizados para obter os resultados foram desenvolvidos utilizando o *software* Matlab. Os dados utilizados nesse capítulo foram coletados utilizando um tempo de amostragem de 10s. Sabe-se que a escolha do tempo de amostragem pode impor limitações em relação às oscilações sendo analisadas, por isso deve-se tomar cuidado na escolha do mesmo. A utilização de um tempo de amostragem grande, por exemplo, pode ocultar oscilações de alta frequência. Neste trabalho, a equipe escolheu o tempo de amostragem e selecionou os dados de modo a visualizar as frequências de interesse para as análises.

O limiar do CV pode variar de acordo com o processo sendo analisado, mas foi fixado para todos os estudos de caso em 0,7%. O limiar de energia utilizado foi de 20%. Os gráficos de espectro serão apresentados utilizando o período em segundos e não frequência em Hz, para uma melhor interpretação dos resultados.

4.1 Pré-tratamento

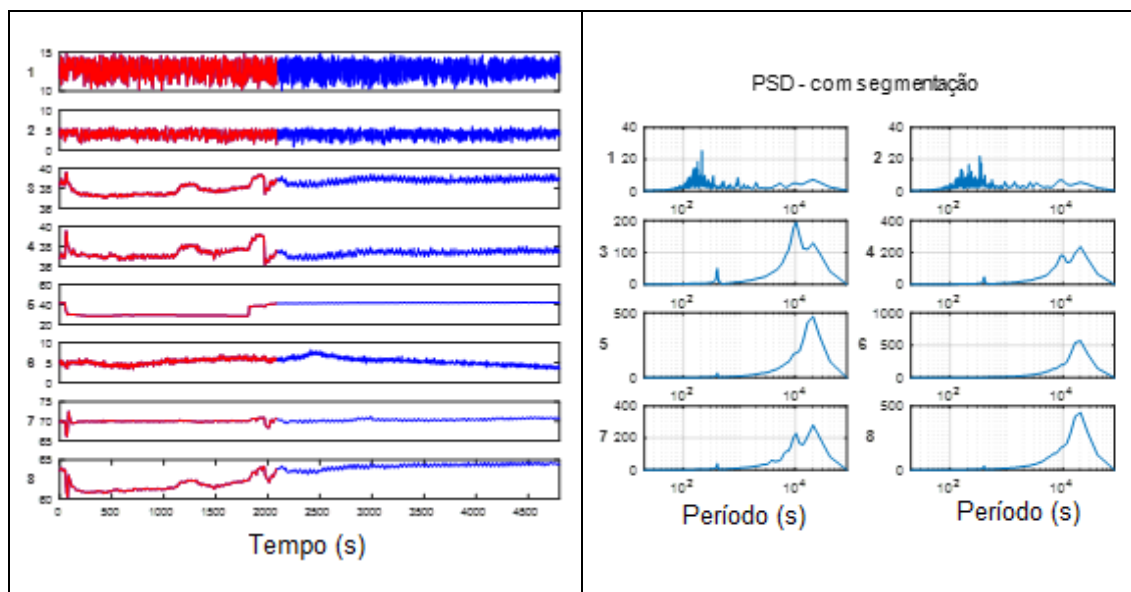
No Capítulo 3, especificamente na Seção 3.1, foram apresentadas três diferentes abordagens para realizar o tratamento nos dados, bem como características de cada tratamento, vantagens e desvantagens. Porém, os 3 tratamentos não devem ser aplicados em conjunto para todos os dados utilizados e também, dos três tratamentos apresentados não existe apenas um que elimina toda a diversidade de distúrbios presentes nos dados. Além disso, um dos objetivos do presente trabalho é automatizar o processo de detecção de oscilações através de um algoritmo. Por esse motivo deve ser escolhido o tratamento que possa ser executado de forma automática para diferentes conjuntos de dados.

Diante dessas restrições, o tipo de tratamento a ser realizado foi definido previamente a partir da análise dos dados, considerando os tipos de distúrbios presentes frequentemente. Uma vez definido, o mesmo tratamento foi utilizado para todos os

estudos de caso deste capítulo. Como a remoção de tendências lineares necessita de seleção manual dos *breakpoints* ou de utilização de um outro algoritmo específico para selecionar os *breakpoints*, optou-se por realizar a segmentação de dados, que resolve o problema com a maioria dos distúrbios não oscilatórios, em conjunto com a filtragem, que resolve o problema com distúrbios oscilatórios de baixa frequência.

O objetivo desta seção é mostrar e discutir os resultados da detecção de oscilações ao se aplicar o conjunto de tratamentos proposto. As análises do Capítulo 3 já mostraram o efeito da utilização dos diferentes tratamentos separadamente, portanto nesta seção o foco é na utilização dos dois tratamentos em conjunto. O conjunto de dados utilizado para esse estudo de caso é mostrado na Figura 24, no gráfico da esquerda. Nota-se que o conjunto de dados possui diferentes tipos de distúrbios. Observa-se em vermelho o segmento de dados extraído pela segmentação e em azul o segmento de dados mantido após o tratamento.

Figura 24 – Resultado da segmentação de dados (em vermelho dados descartados) e PSD após a segmentação de dados

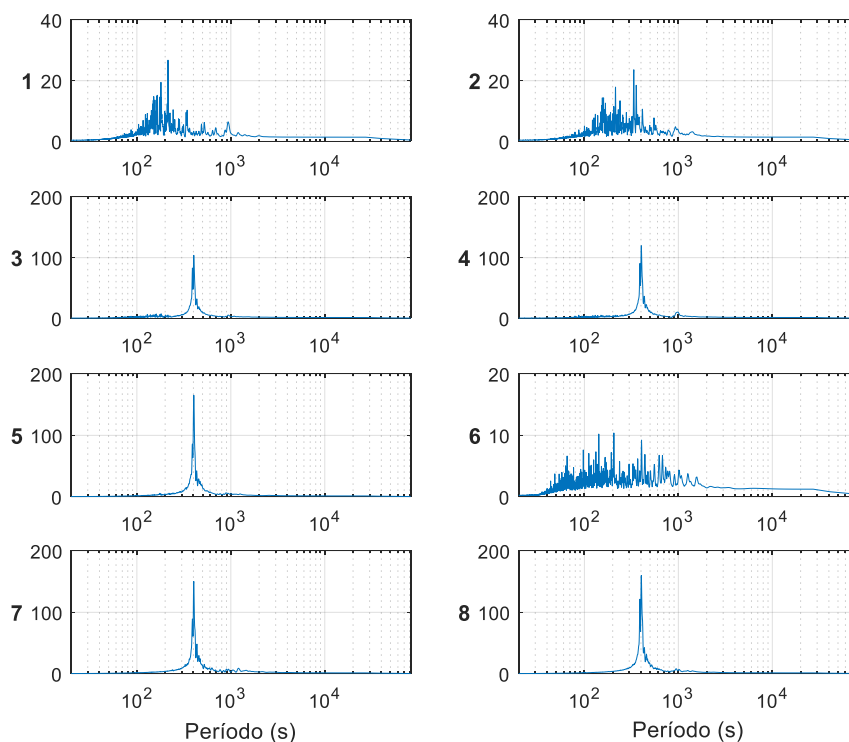


Fonte: O próprio autor.

Uma observação que pode ser feita a respeito dos dados mantidos, em azul, é que os dados não possuem característica estacionária, mesmo após o primeiro tratamento. Além disso, as curvas da direita mostram que distúrbios em baixa frequência ainda estão presentes no espectro. Por esse motivo, em seguida, foi aplicada a filtragem, para lidar com os distúrbios oscilatórios de baixa frequência que prejudicam as análises subsequentes.

Após realizar a filtragem nos dados selecionados pela segmentação de dados, o resultado são as PSD mostradas na Figura 25.

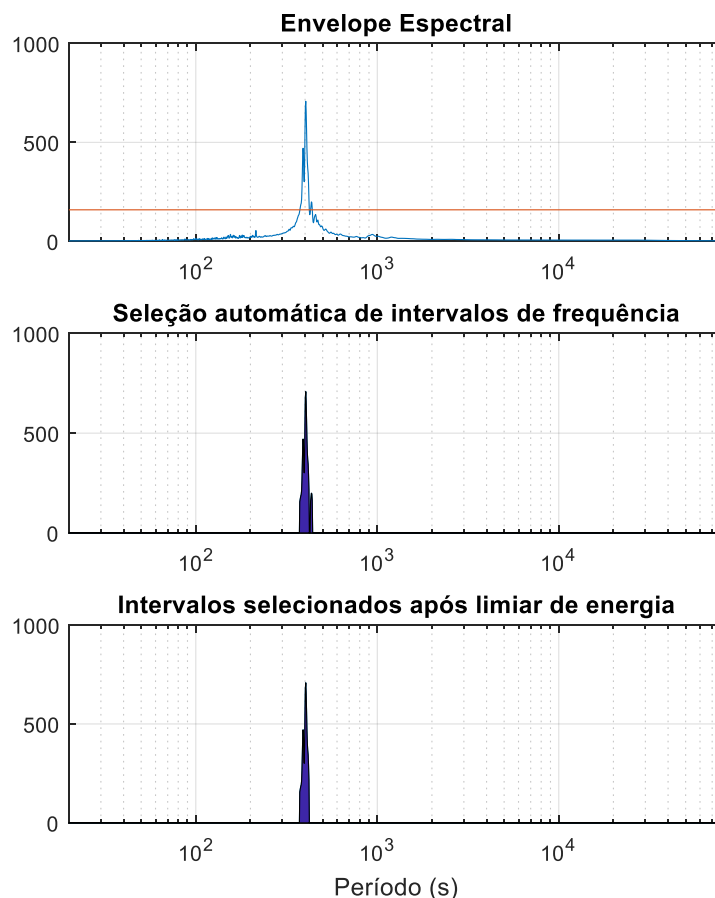
Figura 25 – PSD resultante após a segmentação e a filtragem



Fonte: O próprio autor.

A Figura 26 mostra o envelope espectral calculado utilizando os dados após o tratamento, o resultado da seleção de intervalos de frequência e também o intervalo selecionado após ser aplicado o limiar de energia. Nota-se que dois intervalos foram selecionados sendo que o menor foi descartado pelo limiar de energia.

Figura 26 – Envelope espectral calculado após tratamento (segmentação e filtragem), resultado da seleção automática de intervalos e resultado do descarte de intervalos através do limiar de energia



Fonte: O próprio autor.

A única oscilação selecionada com energia acima do limiar foi uma oscilação de 400s. Seguindo para o resultado da detecção das oscilações utilizando o teste de hipótese, o valor da estatística usada neste teste para a oscilação de 400s nos sinais 3, 4, 5, 7 e 8 foi, respectivamente, 102,5, 74,3, 108, 34,7 e 70,2, acima do limiar de 13,8 para a distribuição qui-quadrado com 2 graus de liberdade. Dessa forma, como a oscilação de 400s foi detectada nos sinais 3, 4, 5, 7 e 8, o resultado está coerente com as oscilações presentes na PSD dos sinais (Figura 25). Já os sinais 1, 2 e 6 apresentaram valores da estatística iguais a 0,4, 0,07 e 1,5, respectivamente.

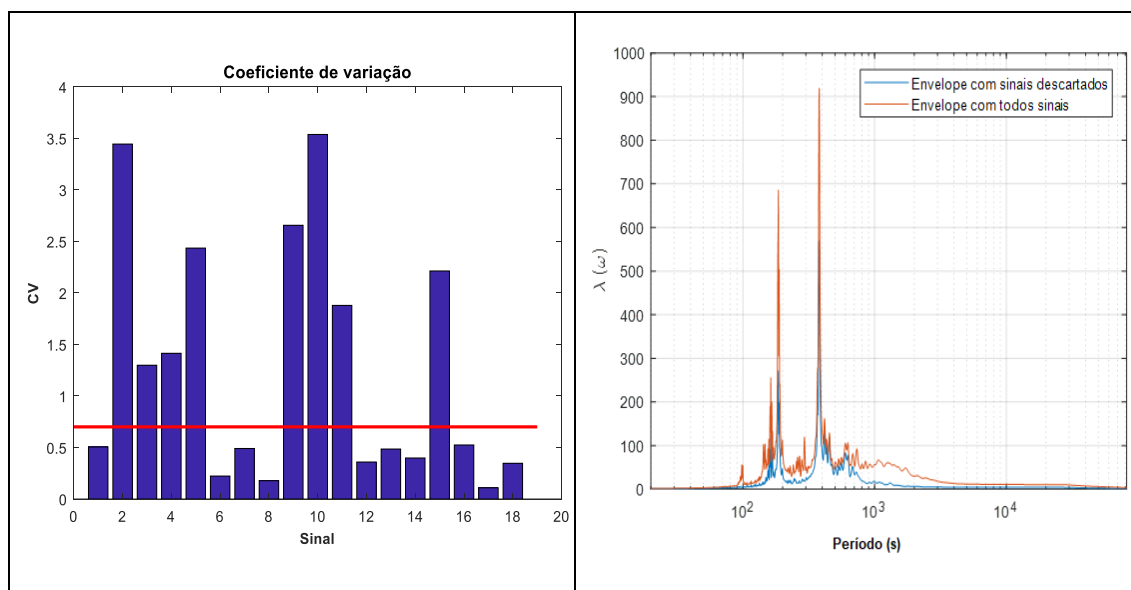
4.2 Impacto das oscilações

Conforme foi mostrado no Capítulo 3, o CV foi a medida de dispersão selecionada para descartar sinais de baixo impacto das análises. Os conjuntos de dados utilizados nesse trabalho podem apresentar grande quantidade de sinais. A decisão de investigar

todos os sinais ou não depende do objetivo que se pretende ao detectar oscilações. Se o objetivo ao investigar o problema é identificar apenas as oscilações de maior impacto, descartar parte dos sinais de baixo impacto é de grande utilidade. Por outro lado, se o objetivo é encontrar causa raiz do problema, oscilação, descartar alguns sinais pode levar ao descarte de um sinal que pode ser a causa raiz. No presente trabalho o objetivo é detectar as oscilações que possuem maior impacto, e não a causa raiz, portanto o descarte de sinais de baixo impacto é necessário.

Utilizando o mesmo conjunto de dados da Seção 3.2, foi feito um comparativo entre as análises utilizando todos os sinais e um outro conjunto de dados menor, contendo apenas os sinais selecionados utilizando o limiar do CV. A Figura 27 mostra, na esquerda, os oito sinais que foram mantidos pelo limiar (2, 3, 4, 5, 9, 10, 11 e 15) e os dez sinais que foram descartados (1, 6, 7, 8, 12, 13, 14, 16, 17 e 18). No gráfico da direita é possível comparar os envelopes calculados com todos os sinais e apenas com os sinais selecionados.

Figura 27 – CV dos sinais analisados e limiar (esquerda) e envelope espectral calculado com todos os sinais e apenas com os sinais selecionados pelo limiar do CV (direita)



Fonte: O próprio autor.

Seguindo análise comparativa, a seleção automática de intervalos de frequência foi utilizada nos dois envelopes calculados, e 2 intervalos de frequência semelhantes foram selecionados (para 185s e 380s), tanto para os sinais selecionados quanto para a análise contendo todos os sinais. A diferença foi notada na detecção de oscilações. Nota-se ao

analisar a Tabela 2 em conjunto com a Figura 27 que as oscilações detectadas quando se utilizaram todos os sinais na análise foram as mesmas, porém, foram encontradas em mais sinais: 7 e 13 para a oscilação 185s e no sinal 8 para a oscilação de 380s, sinais que apresentam baixo CV.

Tabela 2 – Resultado da detecção de oscilações utilizando todos os sinais e apenas os sinais selecionados pelo limiar do CV

Oscilação detectada	Análise com todos sinais	Análise apenas com sinais selecionados pelo CV
185s	Sinais 4, 7, 10 e 13	Sinais 4 e 10
380s	Sinais 3, 8 e 11	Sinais 3 e 11

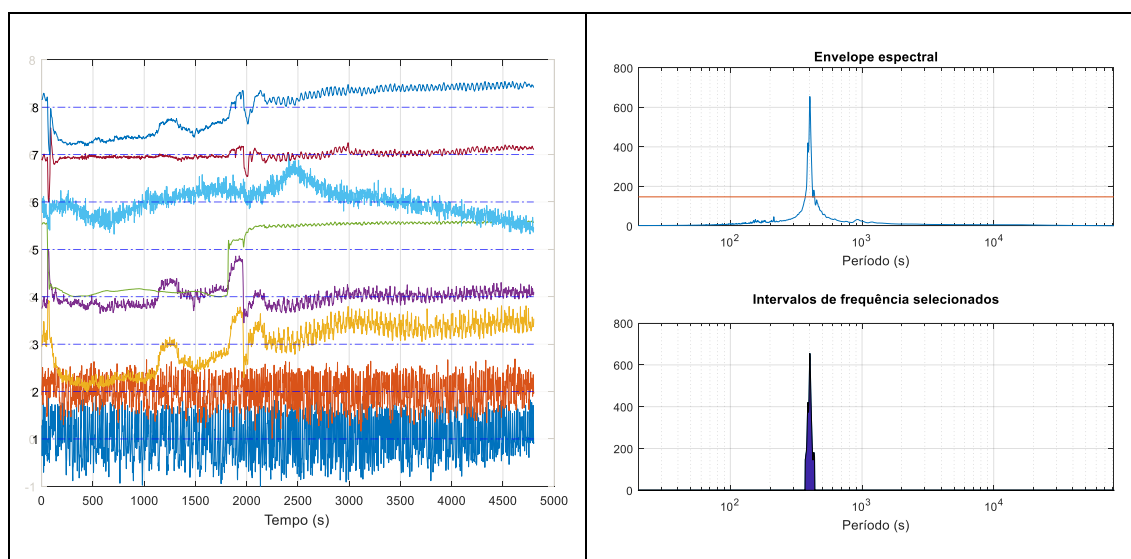
Fonte: O próprio autor.

4.3 Seleção automática de intervalos de frequência

No capítulo 3 foi detalhado o algoritmo utilizado para realizar a seleção automática dos intervalos de frequência. O objetivo da seleção automática é substituir a seleção manual das frequências que devem ser investigadas, conforme foi feito em Jiang et al. (2007).

Nesse capítulo, o objetivo é mostrar o resultado da seleção automática aplicada a diversos conjuntos de dados industriais, e comparar os resultados com a seleção manual das frequências. O primeiro conjunto de dados selecionado é mostrado na Figura 28, à esquerda. O gráfico da direita mostra o envelope espectral e a seleção automática de intervalo de frequência. Neste caso o resultado da seleção foi o mesmo quando comparado a seleção manual (apenas um intervalo de frequência foi selecionado), a única diferença é que foi selecionado um intervalo de frequência no lugar de apenas uma amostra de frequência, um ponto, como foi feito no artigo citado no parágrafo anterior.

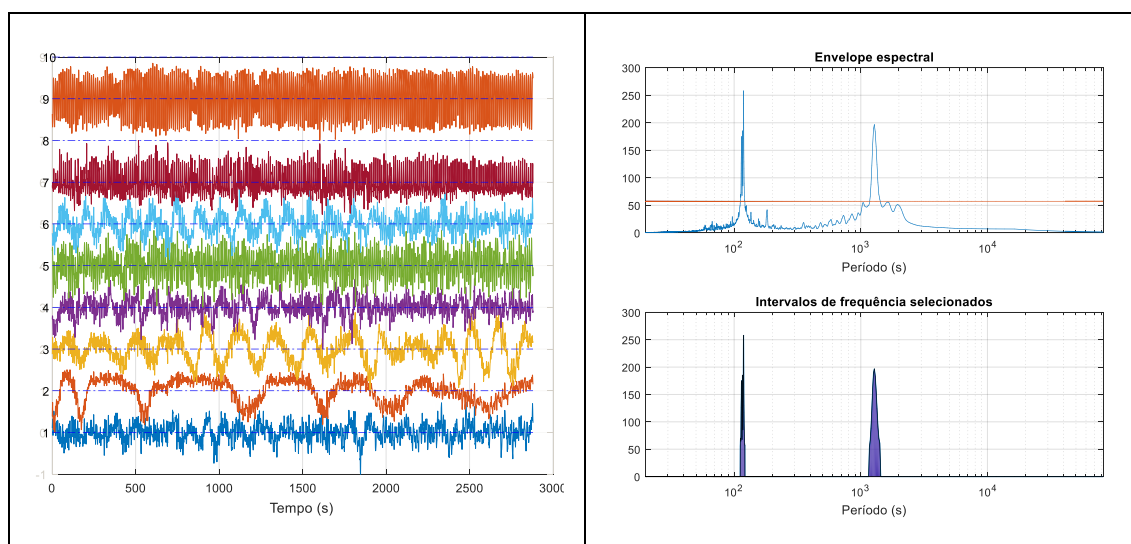
Figura 28 – Comparativo entre a seleção automática e a manual (I)



Fonte: O próprio autor.

Um segundo conjunto de dados foi selecionado (Figura 29) para que fosse feita a comparação entre a seleção automática de intervalos de frequência e a seleção manual. Nesse caso duas oscilações (em 120s e 1280s) foram selecionadas pelo algoritmo e novamente o resultado da seleção manual e automática foram correspondentes.

Figura 29 – Comparativo entre a seleção automática e a manual (II)

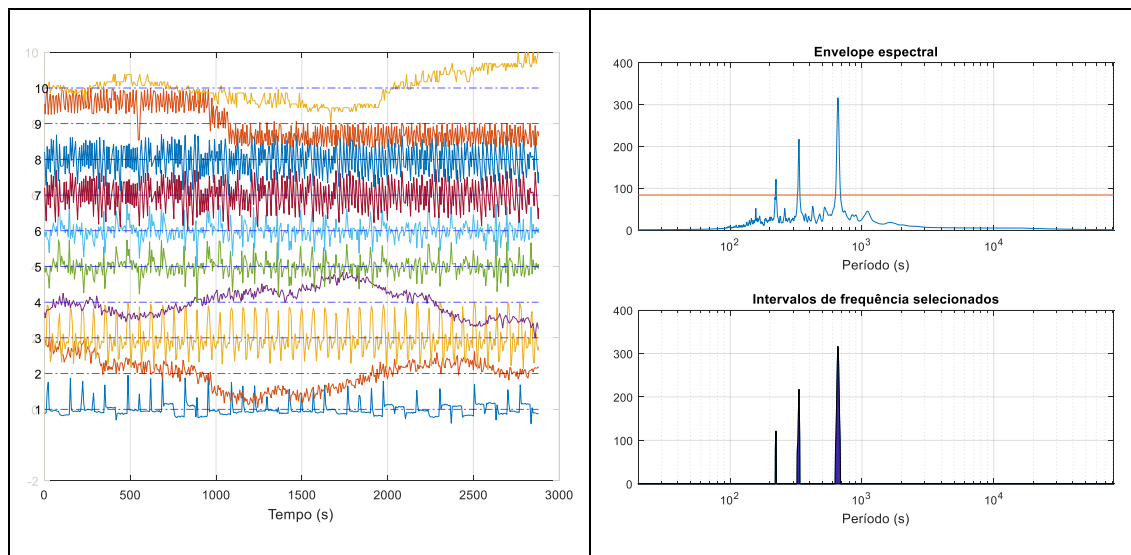


Fonte: O próprio autor.

O terceiro e último conjunto de dados foi selecionado com o propósito de comparar a seleção automática e a seleção manual. O conjunto de dados é mostrado na Figura 30 à esquerda e o envelope espectral e o resultado da seleção automática podem ser vistos no gráfico da direita. Nota-se que três oscilações (220s, 330s e 650s) foram selecionadas

pelo algoritmo e novamente o resultado da seleção manual e automática foram correspondentes.

Figura 30 - Comparativo entre a seleção automática e a manual (III)



Fonte: O próprio autor.

Concluindo esta seção, nos três conjuntos de dados industriais selecionados, verificou-se que a seleção automática foi equivalente à seleção manual. Porém, isso pode não ocorrer em todos os casos, já que situações em que o envelope apresenta diversos picos com mesma amplitude ou amplitude muito próxima podem gerar resultados diferentes da seleção manual de intervalos. A maioria dos estudos de caso mostrou que o algoritmo foi bem-sucedido, e que o procedimento automático pode ser usado.

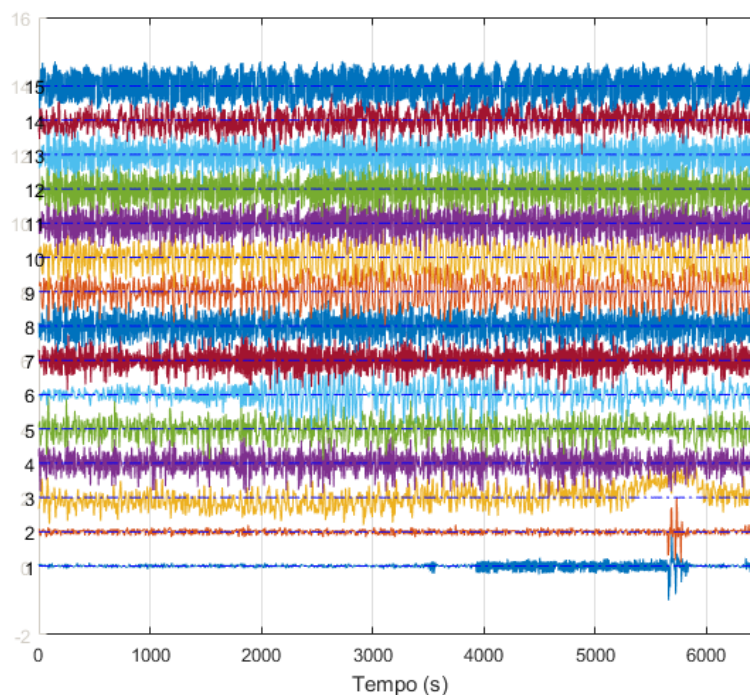
4.4 Repetibilidade

A última seção desse capítulo tem o objetivo de avaliar a característica de repetibilidade dos resultados da metodologia proposta. A repetibilidade será verificada através da detecção de oscilações em intervalos subsequentes de dados.

O conjunto de dados utilizados tem um total de 6480 amostras (18h), dividido em 5 subconjuntos de 1296 amostras cada (3,6h), obtido a partir de sensores de uma unidade de processamento de gás. Em todos os conjuntos de dados foi aplicado o mesmo pré-tratamento: segmentação de dados em conjunto com a filtragem. A escolha foi feita baseada no tipo de distúrbio que estava presente nos 5 subconjuntos de dados utilizados. Também foi utilizado o mesmo limiar do CV igual a 0,7% para todas as análises. O limiar

de energia foi fixado em 20%. O fator de suavização r utilizado foi igual a 3. A Figura 31 mostra os dados de 18h que foram divididos em 5 partes de mesmo tamanho e sequenciais, antes de qualquer tratamento.

Figura 31 – Dados originais sem tratamento (18h)



Fonte: O próprio autor.

O tratamento dos dados foi feito utilizando a filtragem seguindo a metodologia já abordada no capítulo 3. A Tabela 3 mostra o resultado do descarte de sinais pelo CV. Nota-se que na maioria das análises os sinais 1, 2 e 4 foram descartados e em todas delas os sinais 2 e 4 foram descartados.

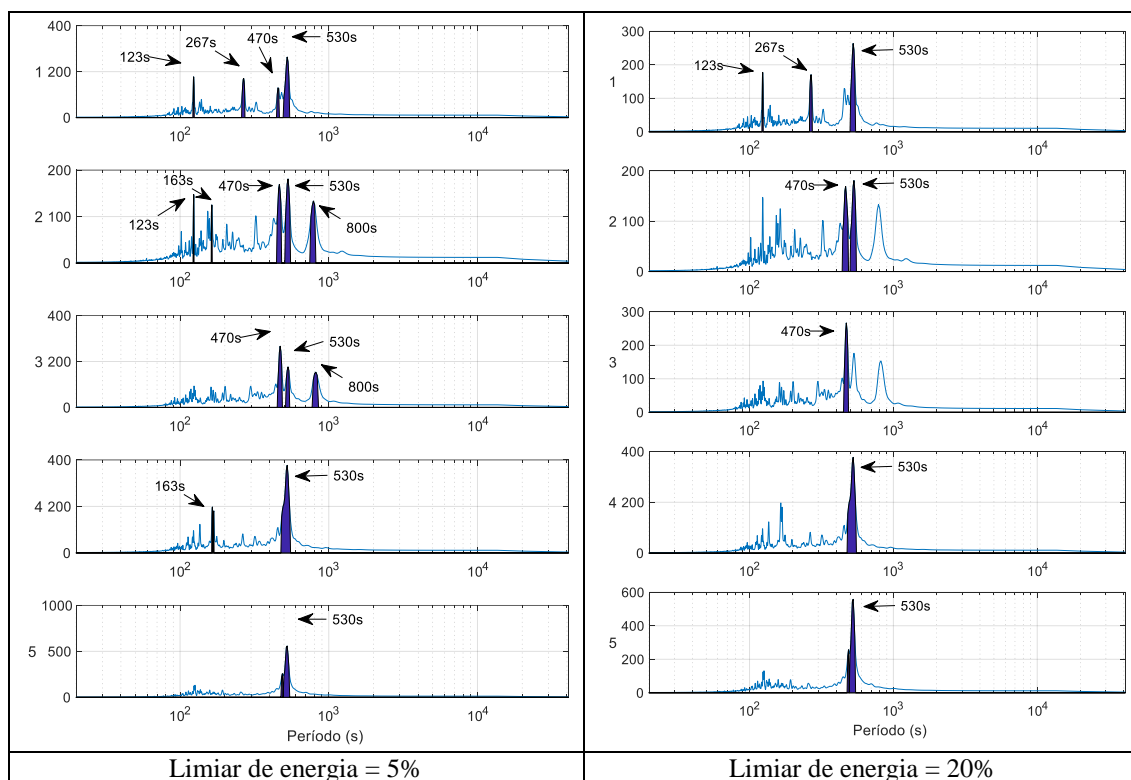
Tabela 3 – Sinais selecionados e descartados pelo limiar do CV nas 5 análises

Subconjuntos/Análise	Sinais mantidos	Sinais descartados
1	3, 5, 8, 9, 10, 11, 12, 13, 14, 15	1, 2, 4, 6, 7
2	3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	1, 2, 4
3	3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	1, 2, 4
4	3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	1, 2, 4
5	1, 3, 5, 6 7, 8, 9, 10, 11, 12, 13, 14, 15	2, 4

Fonte: O próprio autor.

A Figura 32 mostra as 5 curvas do envelope espectral das análises. A área sombreada indica os intervalos de frequência selecionados que apresentaram energia acima do limiar percentual estabelecido. Em cada curva da esquerda foi utilizado um limiar baixo de energia (limiar de 5%), permitindo que muitos intervalos de frequência fossem selecionados. Nas curvas da direita um limiar de energia maior foi utilizado (limiar de 20%), descartando mais intervalos, quando comparadas às curvas da esquerda (limiar de 5%). Nota-se que as oscilações de período 470s e 530s se repetiram em várias análises.

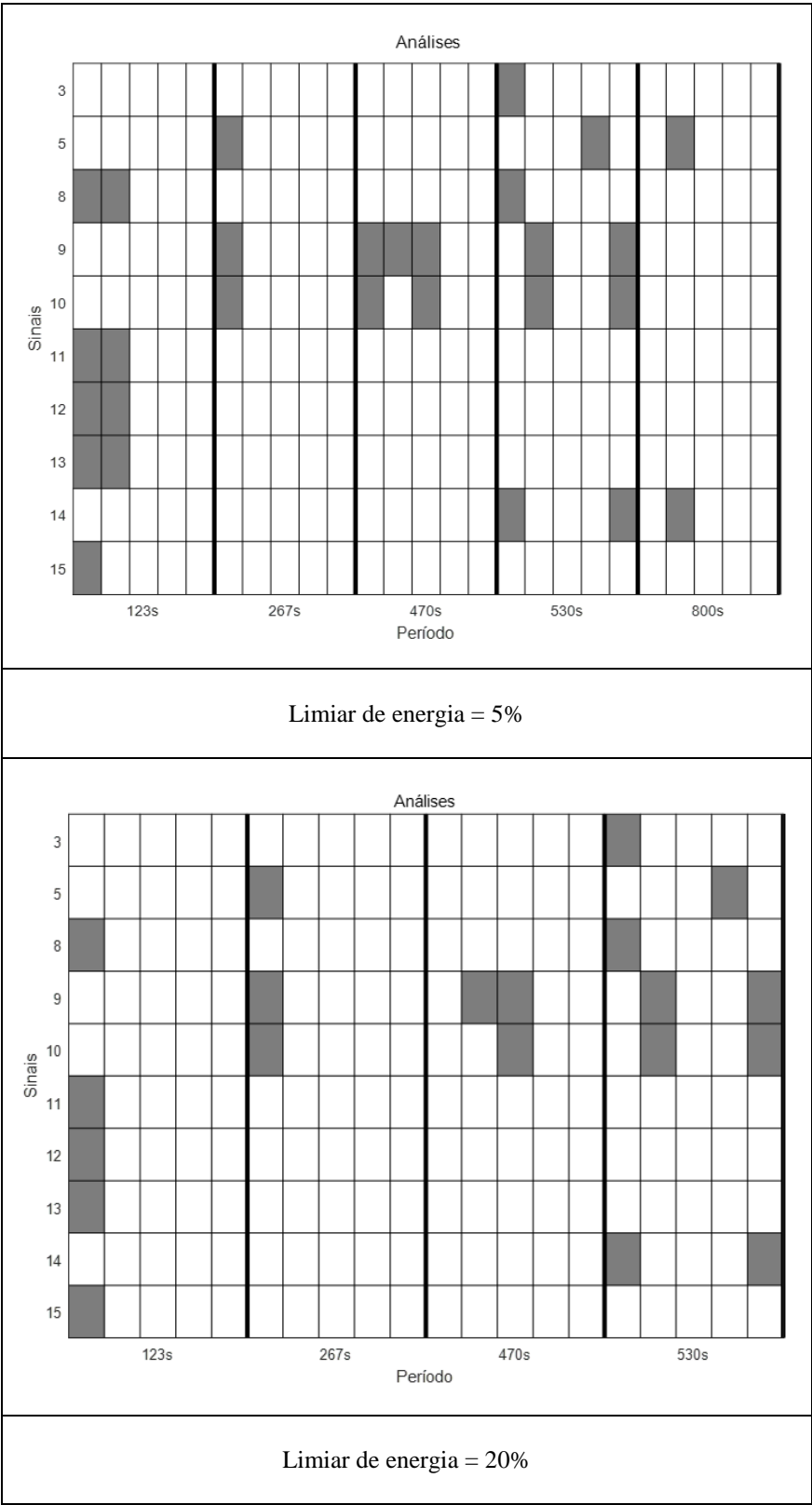
Figura 32 – Envelope espectral e intervalos de frequência selecionados (sombreado azul escuro) nas 5 análises para limiar de energia de 5% e 20%



Fonte: O próprio autor.

Em seguida foi realizado o teste de hipótese para uma amostra dentro dos intervalos de frequência selecionados em cada análise. O resultado das oscilações encontradas nos sinais para todas as 5 análises utilizando os dois limiares de energia pode ser verificado na Figura 33, onde cada retângulo representa uma análise e os preenchidos indicam oscilações detectadas. Nota-se que a oscilação de 530s manteve o maior número de detecções após o limiar de energia ter sido aumentado.

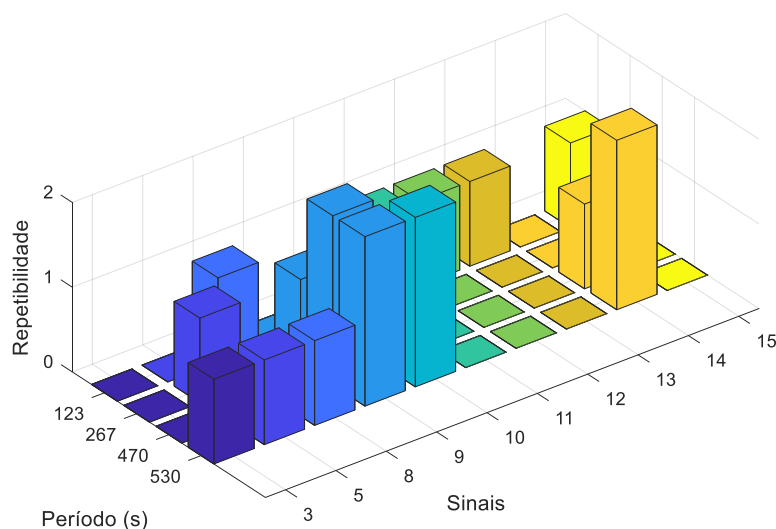
Figura 33 – Resultado da detecção nas 5 análises para limiar de energia de 5% e de 20%



Fonte: O próprio autor.

A Figura 34 mostra o resultado final, para o limiar de energia de 20%, enfatizando a característica de repetibilidade dos resultados de detecção das oscilações após realizar as 5 análises.

Figura 34 – Gráficos de repetibilidade dos resultados para o limiar de energia de 20%



Fonte: O próprio autor.

Conclui-se a partir dos resultados do estudo de caso que as oscilações com período 470s e 530s foram as que mais se repetiram nas 5 análises, sendo que a oscilação de 530s aparece em 4 das 5 análises, indicando há confiança na detecção e confirmando que as oscilações existem e persistem em análises utilizando conjuntos subsequentes. Através da Figura 34 verificou-se que a oscilação de 530s apresentou maior repetibilidade nos sinais 9, 10 e 14 correspondente a sensores de pressão (9 e 10) e nível (14).

O limiar de energia desempenha uma importante função ao descartar intervalos de frequência selecionados que possuem muito ruído. Se os intervalos de frequência estão sendo descartados, essa etapa tem influência direta no resultado da repetibilidade em análises subsequentes, já que a energia de algumas oscilações pode variar dependendo da quantidade de ruído presente no sinal. Este efeito pode ser verificado na Figura 32, onde a oscilação de 800s está presente apenas em algumas análises e quando se aumenta o limiar de energia ela é descartada.

A utilização de um limiar de energia com um valor menor permite que oscilações como a citada no parágrafo anterior estejam presentes em algumas das análises e não se repitam em outras, prejudicando a repetibilidade dos resultados. Porém, as oscilações que

permanecerem ao se aumentar o limiar de energia tendem a se repetir diversas vezes nas análises, conforme foi verificado analisando, em conjunto, os resultados da Figura 32, Figura 33 e Figura 34.

5 Conclusão

Neste trabalho foi apresentada a metodologia do envelope espectral utilizado para detecção de oscilações, abordando os conceitos de estimação do espectro, cálculo do envelope espectral e teste de hipótese para existência da oscilação, que fazem parte da metodologia originalmente proposta na literatura. Foi feita uma análise sobre o efeito da correlação entre os sinais no cálculo do envelope espectral, mostrando sua vantagem sobre a soma dos espectros quando a correlação existe.

Foram propostas contribuições para o método do envelope espectral, entre elas discussões em relação ao pré-tratamento dos dados. Esta é uma importante etapa que deve ser realizada tendo em vista os requisitos para aplicar a metodologia do envelope espectral, sendo o principal requisito a estacionariedade dos dados. Foram abordados três tipos de pré-tratamento: a filtragem, segmentação de dados e a remoção de tendências lineares.

Uma importante contribuição do trabalho foi a utilização do coeficiente de variação para descartar sinais com baixo impacto das análises. A utilização de um limiar para este coeficiente mostrou-se eficiente para descartar tais sinais, permitindo que fossem analisados apenas os sinais relevantes, de modo que a equipe de operação possa concentrar sua atenção nas oscilações de maior impacto.

Foi proposta a utilização de um algoritmo para realizar a seleção automática de intervalos de frequência para identificações das oscilações como substituição da seleção manual empregada na metodologia original, obtendo resultados similares aos obtidos via seleção manual nos estudos de caso apresentados.

Por fim, foi proposta a utilização de um limiar de energia para descartar intervalos de frequência com baixa energia, que contribuiu para aumentar a repetibilidade dos resultados de detecção das oscilações em análises de conjuntos de dados subsequentes.

A metodologia e as contribuições foram aplicadas a dados industriais de unidades de produção de óleo e gás da Petrobrás. Os resultados mostraram que o algoritmo desenvolvido pode ser aplicado de forma automática, produzindo uma boa repetibilidade dos resultados mesmo na presença de distúrbios e ruídos intensos nos sinais analisados.

Referências Bibliográficas

- BIALKOWSKI, W. L. DREAMS VERSUS REALITY - A VIEW FROM BOTH SIDES OF THE GAP. **PULP & PAPER-CANADA**, v. 94, n. 11, p. 19–27, 1993.
- BROWN, R. L.; DURBIN, J; EVANS, J. M. Techniques for Testing the Constancy of Regression Relationships over Time. v. 37, n. 2, p. 149–192, 1975.
- DAMATO, A. **Digital Linear Filters Comparison**. Disponível em: <<https://commons.wikimedia.org/w/index.php?curid=427738>>. Acesso em: 25 jan. 2018.
- DAVIS, R. A.; LEE, T. C. M.; RODRIGUEZ-YAM, G. A. Structural break estimation for nonstationary time series models. **Journal of the American Statistical Association**, v. 101, n. 473, p. 223–239, 2006.
- DE BOOR, C. **A Practical Guide to Splines - Revised Edition**. [s.l], Springer, 2001.
- DURBIN, J. Tests for Serial Correlation in Regression Analysis Based on the Periodogram of Least- Squares Residuals. v. 56, n. 1, p. 1–15, 1969.
- ENDER, B. D. B. Process Control Performance : Not as Good as you Think. **Control Engineering**, v. 40, n. September, p. 180–190, 1993.
- FARIA, H. G. et al. **DETECÇÃO DE OSCILAÇÕES EM PROCESSOS INDUSTRIAIS VIA MÉTODO DO ENVELOPE ESPECTRAL PONDERADOS**. XIII Simposio Brasileiro de Automacao Inteligente. **Anais...**2017. Disponível em: <https://www.ufrgs.br/sbai17/papers/paper_440.pdf>
- GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. [s.l], Amgh Editora, 2011.
- GUO, Z. et al. Automatic detection of multiple oscillations by wavelet analysis. **Computers and Electrical Engineering**, v. 40, n. 7, p. 2167–2177, 2014.
- JIANG, H.; SHOUKAT CHOUDHURY, M. A. A.; SHAH, S. L. Detection and diagnosis of plant-wide oscillations from industrial data using the spectral envelope method. **Journal of Process Control**, v. 17, n. 2, p. 143–155, 2007.
- KARRA, S.; KARIM, M. N. Comprehensive methodology for detection and diagnosis of oscillatory control loops. **Control Engineering Practice**, v. 17, n. 8, p. 939–956, 2009.
- KILLICK, R.; FEARNHEAD, P.; ECKLEY, I. A. Optimal detection of changepoints with a linear computational cost. **Journal of the American Statistical Association**, v. 107, n. 500, p. 1590–1598, 2012.
- KOOPMANS, L. H. **The Spectral Analysis of Time Series**. Second Edi ed. [s.l.] Academic Press, 1995.

- KOTTEGODA, N. T.; ROSSO, R. **Applied Statistics for Civil and Environmental Engineers**. [s.l: s.n.].
- KUCHARCZYK, D.; WYŁOMANŚKA, A.; ZIMROZ, R. Structural break detection method based on the Adaptive Regression Splines technique. **Physica A: Statistical Mechanics and its Applications**, v. 471, p. 499–511, 2017.
- LI, X. et al. The DCT-based oscillation detection method for a single time series. **Journal of Process Control**, v. 20, n. 5, p. 609–617, 2010.
- MATSUO, T.; SASAOKA, H.; YAMASHITA, Y. **Detection and Diagnosis of Oscillations in Process Plants**. Knowledge-Based Intelligent Information and Engineering Systems. **Anais...**2003
- MCDUGALL, A. J.; STOFFER, D. S.; TYLER, D. E. Optimal transformations and the spectral envelope for real-valued time series. **Journal of Statistical Planning and Inference**, v. 57, n. 2, p. 195–214, 1997.
- MIAO, T.; SEBORG, D. E. **Automatic detection of excessively oscillatory feedback control loops**. Proceedings of the 1999 IEEE International Conference on Control Applications (Cat. No.99CH36328). **Anais...**1999Disponível em: <<http://ieeexplore.ieee.org/document/806659/>>
- MITRA, S. K. **Digital Signal Processing - Computer Based Approach - Sanjit K. Mitra.pdf****Microelectronics Journal**, 2001. Disponível em: <http://www-ee.eng.hawaii.edu/~treed/EE415/DSP_Errata.pdf>
- RABINER, L. R.; GOLD, B. **Theory and application of digital signal processing**. [s.l: s.n.].
- SHINSKEY, F. G. How good are our controllers in absolute performance and robustness? **Measurement and Control**, v. 23, n. 4, p. 114–121, 1990.
- SHUMWAY, R. H.; STOFFER, D. S. Time series analysis and applications. 2006.
- SHUMWAY, R. H.; STOFFER, D. S. **Time Series Analysis and Its Applications With R Examples**. [s.l.], Springer, 2011.
- SMITH, S. W. **The Scientist and Engineer's Guide to Digital Signal Processing**. Second ed. [s.l.] California Technical Publishing, 1999.
- SRINIVASAN, B.; RENGASWAMY, R. Automatic oscillation detection and characterization in closed-loop systems. **Control Engineering Practice**, v. 20, n. 8, p. 733–746, 2012.
- SRINIVASAN, R.; RENGASWAMY, R.; MILLER, R. A modified empirical mode decomposition (EMD) process for oscillation characterization in control loops. **Control**

Engineering Practice, v. 15, n. 9, p. 1135–1148, 2007.

STOFFER, D. S. Detecting Common Signals in Multiple Time Series Using the Spectral Envelope. **Journal of the American Statistical Association**, v. 94, n. 448, p. 1341–1356, 1999.

STOFFER, D. S.; TYLER, D. E.; MCDUGALL, A. J. Spectral analysis for categorical time series: Scaling and the spectral envelope. **Biometrika**, v. 80, n. 3, p. 611–622, 1993.

STOICA, PETRE; MOSES, R. **Spectral Analysis of Signals**. [s.l.] Prentice Hall, 2005.

THORNHILL, N. F. et al. Spectral principal component analysis of dynamic process data. **Control Engineering Practice**, v. 10, n. 8, p. 833–846, 2002.

THORNHILL, N. F.; COX, J. W.; PAULONIS, M. A. Diagnosis of plant-wide oscillation through data-driven analysis and process understanding. **Control Engineering Practice**, v. 11, n. 12, p. 1481–1490, 2003.

THORNHILL, N. F.; HUANG, B.; ZHANG, H. Detection of multiple oscillations in control loops. **Journal of Process Control**, v. 13, n. 1, p. 91–100, 2003.

VIEIRA JUNIOR, R. **COMPARAÇÃO DE MÉTODOS PARA DETECÇÃO DE OSCILAÇÕES EM UNIDADES INDUSTRIAIS AFETADAS POR DISTÚRBIOS**. [s.l.] UFES, 2016.

WANG, J.; HUANG, B.; LU, S. Improved DCT-based method for online detection of oscillations in univariate time series. **Control Engineering Practice**, v. 21, n. 5, p. 622–630, 2013.

WENDT, D. A.; TYLER, D. E.; STOFFER, D. S. The spectral envelope and its applications. **Statistical Science**, v. 15, n. 3, p. 224–253, 2000.

XIA, C.; HOWELL, J. Isolating multiple sources of plant-wide oscillations via independent component analysis. **Control Engineering Practice**, v. 13, n. 8, p. 1027–1035, 2005.

XIA, C.; HOWELL, J.; THORNHILL, N. F. Detecting and isolating multiple plant-wide oscillations via spectral independent component analysis. **Automatica**, v. 41, n. 12, p. 2067–2075, 2005.

XIE, L. et al. Online oscillation detection in the presence of signal intermittency. **Control Engineering Practice**, v. 55, p. 91–100, 2016.

6 APÊNDICE – Rotinas e códigos desenvolvidos em Matlab

A maioria das rotinas foi desenvolvida pelo autor, utilizando a menor quantidade possível de funções do Matlab. Por exemplo, a estimação do espectro utilizando funções de correlação, que poderia ter sido feita utilizando funções que geram o periodograma e o correlograma. Essa escolha foi feita com o objetivo de manter os cálculos mais próximos da metodologia do envelope espectral descrita em Jiang et al.; (2007), tendo assim um maior controle sobre cada etapa da metodologia. Porém, para casos em que a metodologia é bem definida em diversos livros e trabalhos, como por exemplo a utilização do filtro de Butterworth, optou-se por utilizar funções do próprio Matlab.

O algoritmo de detecção de oscilações utilizando o envelope espectral, descrito no Capítulo 3, foi implementado utilizando o software Matlab. Foi desenvolvida uma rotina principal, que realiza o tratamento dos dados, seleção utilizando o CV, pré-tratamento dos dados e também responsável por gerar os gráficos que mostram os resultados parciais de cada etapa e também o resultado final da detecção de oscilações. Também foram desenvolvidas funções auxiliares com objetivos específicos, visando uma melhor organização do algoritmo de forma geral, bem como a utilização das funções para análises específicas.

O algoritmo principal possui alguns parâmetros que devem ser escolhidos pelos pesquisadores ou pela equipe de operação, sendo eles: tempo de amostragem (em segundos), fator de suavização, limiar de energia, limiar do CV e os parâmetros da rotina de segmentação de dados, já detalhados no Capítulo 3.

6.1 Cálculo do CV e descarte de sinais com baixo impacto

```
mediad = mean(Xm1,2);  
stdd = std(Xm1,0,2);  
  
medx = mediad;  
stdx = stdd;  
  
test = find(medx < 1);  
medx(test) = 1;  
CV = 100*stdx./medx;  
Sid = find(CV > limiar_cv);
```

6.2 Normalização, filtragem e segmentação de dados

```
mediad1 = mean(Xm1,2);  
stdd1 = std(Xm1,0,2);  
  
for i = 1:N  
    Xnn(i,:) = (Xm1(i,:)-mediad1(i))./stdd1(i);  
  
end  
  
%% Filtro  
f1 = 1/1200;  
fs = 1/ts;  
ordemfilt = 2;  
  
for i=1:N  
    [b1,a1] = butter(ordemfilt, f1/(fs/2), 'high');  
    saida = filter(b1,a1,Xm1(i,:));  
  
    X_s1(i,:) = saida;  
  
end
```

```

function ind1=select2(x,nb,ndp,nd,limiar)
if nargin==1
    ndp=5;
    nd=3;
    nb=50;
    limiar=0.025;
end;
I=[];
N=length(x);
[n,cb]=hist(x,nb);
nn=100*n/N;

[nm,ii]=max(nn);
I=ii;k=1;
if ii<nb
    while nn(ii+k)>limiar*nn(ii)
        I=[I ii+k];
        k=k+1;
        if (ii+k)>nb break; end;
    end;
end;
k=1;
if (ii>1)
    while nn(ii-k)>limiar*nn(ii)
        I=[I ii-k];
        k=k+1;
        if (ii-k)==0 break; end;
    end;
end;
l1=min(I);l2=max(I);
if (l1-1)==0 l1m=min(x); else l1m=cb(l1-1); end;
if (l2+1)>nb l2m=max(x); else l2m=cb(l2+1); end;
l1=mean([l1m cb(l1)]);
l2=mean([cb(l2) l2m]);
ind=find((x>l1)&(x<l2));
ind=unisel(x,ind,ndp);
ind1=segselect2(ind,nd);
if nargin==0 plot(1:N,x,ind1,x(ind1),'r*'); end;
end
%
```

```

function ind1=unisel(x,ind,ndp)
m1=mean(x(ind));
s1=ndp*std(x(ind));
z=abs(x-m1);
ind1=find(z<s1);
end
%
```

```

function Tind=segselect2(Ind,nd)
df=diff(Ind);
Nd=find(df>nd); % Pontos nos quais há descontinuidade de pelo menos nd
amostras
if length(Nd)==0
    Tind=Ind;
else
    [L,C]=size(Nd);
    if C<L Nd=Nd'; end;
    Ib=[1 Nd length(Ind)];
    dIb=diff(Ib);
    [~,I Ib]=max(dIb); % Encontra maior segmento possível
    %Tind=[Ind((Ib(I Ib)+50):(Ib(I Ib+1)-1))]; %Seleciona indices do segmento
    (Ib(I Ib)+50 para eliminar ainda efeitos de outliers)
    Tind=[Ind((Ib(I Ib)):(Ib(I Ib+1)))];
end;
end
```

6.3 Função para estimação do espectro e cálculo do envelope espectral

```
function
[CorrMf,M2,B,Beta,Lamb,wk,indicevet,lambt,Betat]=envelope_corr(X,ts,smo)
% Programa de Pós-Graduação em Engenharia Elétrica - 27/08/2017
% Envelope_corr: Funcao responsavel por calcular o envelope espectral
% utilizando o correlograma

[L,N]=size(X);

NFFT = 2^nextpow2(2*L);

wk = (1:1:floor(NFFT/2))./(NFFT);

wk = wk';

Mvet=[];F=[];
vet=exp(-2*pi*wk*sqrt(-1));

for i=1:NFFT/2
% for i=1:L
    Mvet=[Mvet vet.^(i-1)];
end;

for i=1:N
    for j=1:N
        [c,l]=xcorr(X(:,i),X(:,j),NFFT/2,'Coeff');

        c=c(NFFT/2:(NFFT-1))';

        CorrMult=abs(c*Mvet);

        CorrMultf=filter(ones(1,smo)/smo,1,CorrMult); % Suavização

        CorrMf(i,j)={CorrMultf}; % Espectro (Px)
    end;
end;

M2={};
for i=1:N
    for j=1:N
        for k=1:length(wk)
            M2{k}{i,j}=CorrMf{i,j}(k);
        end
    end
end

Lamb=[];
B=[];
Beta = [];
indicevet = [];
vx = cov(X);

% Calcular o envelope atraves do maior autovalor de Px
for i=1:length(wk)
    avet = [];
    aval = [];
    avet2 = [];
    aval2 = [];
```



```

[avet2,aval2] = eig(cell2mat(M2{i}));

indice = 0;

[valor indice]=max(max(aval2));
lambt{i} = aval2;
indicevet = [indicevet indice];

Lamb(1,i)=aval2(indice);

Bnew = avet2;
Beta=[Beta Bnew(:,indice)];

Betat{i} = avet2;

end

% disp('teste')
wk = 1/ts.*wk;

end

```

6.4 Função para seleção automática dos intervalos de frequência e descarte de intervalos utilizando limiar de energia

```
function [Intf,Intflim,limiart,energy_int,envelopefinal] =
energy_interval(envelope,wl,penerg)
% Programa de Pós-Graduação em Engenharia Elétrica - 27/08/2017
% Realizar uma varredura descendente a partir do maior pico do envelope com
% o objetivo de selecionar os principais intervalos de frequência do
% envelope. Em seguida calcula a energia percentual desses picos e retorna
% apenas os intervalos com energia maior que o limiar.

%% STEP 1: Intervalos do envelope
Intf = [];
Intflim = [];

% Trecho do código para definir o limiar
[pk2,indpk] = findpeaks(envelope);
[pk2m indd]= max(pk2);
indpkm = indpk(indd);

antienvelope = -envelope;

[pk3,indpk3] = findpeaks(antienvelope);
[pk3m,indd3]= max(pk3);
indpkm3 = indpk3(indd3);

amp = pk2m - (abs(pk3m));

limit = 6;
lold = 0;
limiart = pk2m; % Ponto de partida da busca (Limiar)
% divamp = amp/5;
% divamp = amp/4.5;
divamp = amp/4.75;

% Loop varredura limiar
while(1)
    limiart = limiart - divamp;
    indt = crossing(envelope,wl,limiart);
    lindt = length(indt)/2;
    limit2 = lindt - lold;

    if (limit2 > limit || limiart <= 0)
        limiart = limiart + 1.5*divamp; % TESTE
        break;
    else

        lold = lindt;
    end
end

% Encontrar intervalos de frequência para o limiar
ind3=crossing(envelope,wl,limiart);
if (ind3(1) == 1)
    ind3(1) = [];
elseif (ind3(end) == length(ind3))
    ind3(end) = [];
end
```

```

if isempty(ind3)
    return;
else
    ind3=ind3+1; %
end;

%% STEP 2: Plot teste avaliar limiar
figure
testlinha = limiart*ones(1,length(1./w1));
semilogx(1./w1,envelope,1./w1,testlinha)
grid on
xlabel('Período (s)');
ylabel('\lambda (\omega)');

%% Corrigir numero impar de cruzamentos por zero
vetdiff = [];
for i=1:size(ind3,2)
    vetdiff = [vetdiff ;ind3(i)-1 ind3(i)+1];
end

teste = envelope(vetdiff);

difer = diff(teste,1,2);

if(difer(1) < 0)
    ind3(1) = [];

elseif (difer(end)>0)
    ind3(end) = [];
end

l=length(ind3);

M=l/2; % M possíveis faixas de frequências presentes no sinal
wil=[];

if M ==0
    disp('Erro: Nenhum intervalo encontrado!! ')
    return;
end

% Obtém os valores de frequência em que os cruzamentos ocorreram
for i=1:l
    wil(1,i)=w1(ind3(i)); % Intervalos de frequência obtidos do PSD
end;

Intf = ind3; % Indices dos intervalos de frequencias selecionadas
Intwi = wil; % Intervalos de frequencias selecionadas

%% STEP 5: Verific. de energia dos intervalos do envelope
ind = Intf;
wi = wil;

M=l/2; % M possíveis faixas de frequências presentes no sinal

```

```

wiaux=[];

%Avalia se os intervalos de frequencia detectados
bandas=[];
energy = [];

% Energia total do envelope
for i = 1:M

energy_full(i) = sum(envelope(ind(2*i-1):ind(2*i)));

end
energy_int = [];
energyfull2 = sum(envelope);
energyt_full = energyfull2;
total_energy = energyt_full;
limiar_t = penerg*energyt_full;

for i=1:M
    Energyi(i)=sum(envelope(ind(2*i-1):ind(2*i)));
    Energyiperc(i) = (Energyi(i)/energyt_full)*100;
    energy_int = [energy_int;Energyiperc(i)];
    if Energyi(i)>=limiar_t
        energy=[energy;Energyi(i)];
        wiaux=[wiaux [wi(2*i-1) wi(2*i)]]; % mantém apenas faixas com Energia
    > penerg
        bandas=[bandas; [ind(2*i-1),ind(2*i)]];
    end;
end;
energy=energy*100;
wi2=wiaux;
M=size(wi,2)/2; % M Frequencias presentes no sinal com mais de p% de
energia

if M==0
    return;
end;

indb = [];
for i = 1:size(bandas,1)
    for j = 1:size(bandas,2)
        indb = [indb bandas(i,j)];
    end
end

end

Intflim = indb; %
Intwi = wi2;

envelopefinal = envelope;

% FIM
return;
end

```

6.5 Função para realizar a indicação de sinais que possuem oscilações utilizando teste de hipótese

```
function [F,Mresult ] =
hypo_test(Betat2,lambt2,Lamb2,Beta2,r,wk2,Intflim,X,Sid)
%% STEP 10: Teste de hipotese
for i = 1:length(Betat2)

    Betat2{i} = abs(Betat2{i});
    lambt2{i} = abs(lambt2{i});

end

h = [];
for j = -r:r
    h1 = (r-abs(j)+1)/((r+1).^2);
    h = [h h1];
end

vn = 0;
for j = 1:length(h)

    vn = vn+(h(j).^2);

end
vn = vn.^0.5;
preprod = (vn.^(-2))*Lamb2;

for i = 1:size(Beta2,2)
    premat = [];
    premats = [];
    for j = 2:size(Beta2,1)

premat = (lambt2{i}(j,j)*(lambt2{i}(1,1)-lambt2{i}(j,j)).^(-
2)*Betat2{i}(:,j)*conj(transp(Betat2{i}(:,j)))));

        if j == 2
            premats = premat;
        else
            premats = premats + premat;
        end

    end
    Vb{i} = preprod(i)*premat;

end

% LEMBRAR DE USAR APENAS A DIAGONAL DE VB PARA CADA FREQUENCIA (SIGMA)
sigmab = [];
for i = 1:length(Vb)
    sig = diag(Vb{i});
    sigmab = [sigmab sig];
end

% NO TESTE DE HIPOTESE USAR APENAS O BETA1 OU SEJA, O MAIOR AUTOVETOR
hypemat = [];
hypemat2 = [];
for i = 1: size(Beta2,2)
    hypet = (2./sigmab(:,i)).*(abs(Beta2(:,i)).^2);
    hypet2 = (2.*(abs(Beta2(:,i)).^2))./sigmab(:,i);
```

```

    hypemat = [hypemat hypet];
    hypemat2 = [hypemat2 hypet2];
end
hypemat = [];
hypemat = hypemat2;

achar = (hypemat == Inf);
hypemat(achar) = 0;
hypemat(:,1) = 0;

% Plot do teste de hipotese
figure
limiteh = chi2inv(0.999,2);
limitehv = limiteh*ones(1,length(hypemat));

for i = 1:size(hypemat,1)
    if(size(hypemat,1) <= 3)
        subplot(size(hypemat,1),1,i)
    else
        if(mod(size(hypemat,1),2)==1 && i == size(hypemat,1))
            subplot(ceil(size(hypemat,1)/2),2,[i i+1])
        else
            subplot(ceil(size(hypemat,1)/2),2,i)
        end
    end

    semilogx(1./wk2,hypemat(i,:),1./wk2,limitehv)

    ylabel(num2str(Sid(i)), 'HorizontalAlignment', 'right', 'VerticalAlignment', 'middle', 'FontWeight', 'bold')
    set(get(gca, 'YLabel'), 'Rotation', 0)
    vett = 1./wk2;
    xlim([vett(end) vett(1)])
    lowerl = limiteh - 30;
    upperl = limiteh + 30;
    ylim([lowerl upperl]);
    xlabel('Frequency(s)', 'FontSize', 10, 'FontWeight', 'bold')

end
suptitle('Hypothesis test results')
%% STEP 11: Indicar sinais onde existem as oscilações utilizando teste de hipotese

ind_full = zeros(1,size(X,1));
ind_full(Sid) = Sid;

newhypemat = zeros(length(ind_full),size(hypemat,2));
newhypemat(Sid,:) = hypemat;

hypemat = [];
hypemat = newhypemat;

k = 0;
linhasmat = length(Intflim)/2; %frequencias detectadas
colunasm = size(hypemat,1); % sinais
matrizresult = zeros(linhasmat,colunasm);
for i = 1:2:length(Intflim)
    k = k+1;
    mediafreq(k) = mean(wk2(Intflim(i):Intflim(i+1)));
    for j = 1:size(hypemat,1)
        interval1 = hypemat(j,Intflim(i):Intflim(i+1));
    end
end

```

```
resultp = find(intervall > limiteh);  
resultp = resultp + Intflim(i) - 1;  
    if size(resultp,2) >= 1  
        matrizresult(k,j) = 1/mediafreq(k);  
    end  
  
end  
  
end  
  
Mresult = matrizresult;  
F = mediafreq;  
  
end
```